

Département de Géographie et Aménagement 2001 / 2002
Institut Daniel Faucher
Atelier de Cartographie - Infographie
Laurent JÉGOU et Joseph BUOSI

INITIATION
A LA CARTOGRAPHIE AUTOMATIQUE

Avant propos

Ce cours vise à présenter les méthodes et techniques de la cartographie statistique par ordinateur, dite cartographie automatique. Il se base sur les acquis du cours de cartographie générale, pour présenter les avantages de l'automatisation grâce à l'outil micro-informatique.

Ce cours est disponible en format informatique (PDF) à l'adresse suivante : http://www.univ-tlse2.fr/geoprdc/scap

L'essentiel des méthodologies présentées requiert l'utilisation d'un logiciel de cartographie automatique performant, accompagné si possible d'un (logiciel) tableur. La bonne compréhension de ce cours nécessite donc une importante phase de travaux personnels, de pratique des méthodologies, c'est pourquoi chaque leçon est complétée par la manipulation d'un tel logiciel en situation réelle sous la forme d'exercices guidés.

Comme la disponibilité d'un logiciel adapté n'est pas très courante, l'Atelier de Cartographie de l'UTM a réalisé un logiciel spécifiquement orienté vers l'apprentissage de la cartographie automatique. Ce logiciel, SCAP, est disponible gratuitement dans une version limitée à un seul fond de carte (celui utilisé dans le présent cours), soit sur le support Cd-rom de ce cours soit sur le site de SCAP : http://www.univ-tlse2.fr/geoprdc/scap. Vous pouvez aussi vous le procurer en version complète au prix moyen d'un livre de cours auprès de l'Atelier (cf. la page "contact" dans le chapitre "compléments"). Ce logiciel n'est pas utilisé explicitement dans le cours mais il constitue un complément idéal, car il propose une automatisation de toutes les techniques présentées.

Bibliographie :

BEGUIN Michèle, PUMAIN Denise - 1994 - La représentation des données géographiques- Ed. Armand Colin - Paris.

BERTIN Jacques - 1977 - La Graphique et le Traitement Graphique de l'Information. Ed. Flammarion - Paris.

BLIN Eric, BORD Jean-Paul - 1993 - Initiation géo-graphique - Ed. Sedes – Paris

BONIN Serge - 1975 - Initiation à la graphique - Ed. Epi – Paris

BRUNET Roger - 1987 - La Carte Mode d'Emploi. - Ed. Fayard/RECLUS – Paris

SERVICE TECHNIQUE DE L'URBANISME - 1989 - Chiffres et Cartes : une union réfléchie. Ministère de l'Équipement et du Logement / RECLUS.

CAUVIN Colette, REYMOND Henri, SERRADJ Abdelaziz - 1987 " Discrétisation et représentation cartographique." - Ed. RECLUS, Montpellier.

Didier POIDEVIN, « La Carte, moyen d'action », Ellipses, Paris, 1999.

En anglais :

TUFTE Edward R. – 1983 – The Visual Display of Quantitative Information, Graphics Press.

TUFTE Edward R. – 1997 - Visual Explanations, Images and Quantities, Evidence and Narrative.

CARTOGRAPHIE AUTOMATIQUE : POSITIONNEMENT ET PRÉ REQUIS

Le présent cours développe les techniques de la cartographie générale dans un domaine particulier : la conception cartographique assistée par micro-ordinateur.

La CCAO va aider le cartographe dans les différentes phases de la réalisation d'une représentation cartographique, de l'analyse/traitement de la variable à l'édition du document final, mais *elle ne s'applique que dans certains cas et qu'à certains types de carte.*

Ce cours va donc reprendre une grande partie du vocabulaire et des acquis du cours de cartographie générale, qu'il est donc préférable d'avoir parcouru auparavant. On reprendra en effet tous les concepts concernant les variables graphiques, l'imposition et la recherche des composantes, tout en maintenant les mêmes objectifs généraux de lisibilité, d'objectivité et d'efficacité.

On va donc ici développer une petite partie des techniques de cartographie générale pour présenter les méthodes et les outils qui permettent de les automatiser. Grâce à ces techniques, la réalisation complète d'une carte thématique est automatisée dans de nombreux aspects, notamment dans ses opérations les plus longues et fastidieuses, pour permettre au cartographe d'aboutir rapidement à une représentation graphique d'une valeur statistique sur un fond de carte.

Mais cette rapidité potentielle d'exécution de la carte peut être perçue comme une invitation à laisser le logiciel décider à votre place. Ce serait utiliser un outil sans en guider les mouvements, ce qui aboutit souvent à des accidents...

La CCAO est donc un ensemble de techniques dont il faut maîtriser les fonctions et le contexte correctement avant de pouvoir les utiliser, sous peine de réaliser, très rapidement certes, des cartes, mais des cartes fausses ou peu lisibles. C'est pourquoi, avant d'utiliser de tels outils, il faut en comprendre le fonctionnement et l'utilité précise. En conclusion, voici les principes qui vont guider le présent cours.

Ce cours de cartographie automatique est structuré par les principes suivants :

- La carte à réaliser doit être la représentation la plus claire et la plus fidèle (objective) de l'information donnée.
- La carte à réaliser sera thématique et non synthétique.
- La variable à représenter est quantitative, donc ordonnée et proportionnelle.

DÉTERMINATION DE L'OBJECTIF DE LA REPRÉSENTATION CARTOGRAPHIQUE

Toute représentation est une simplification.

Une représentation graphique, par le biais de formes dessinées, suppose toujours une simplification importante de l'information d'origine. C'est naturellement le cas lorsque l'on passe d'une information numérique, sous forme de séries de chiffres, à une position sur un graphique, sur une carte.

Une variable quantitative devient alors une position sur un plan à deux dimensions, et dans le cas de la carte, ces deux dimensions sont déjà occupées par le fond de carte...

Mais c'est le principe même de la représentation graphique : une fois simplifiée et représentée, l'information est beaucoup plus rapidement assimilable, compréhensible. Au lieu d'avoir à lire plusieurs colonnes de chiffres, le lecteur n'aura qu'un seul graphique ou qu'une seule carte à observer.

- *Représenter c'est simplifier*
- *Simplifier c'est réduire l'information*
- *La représentation graphique doit préserver au mieux l'information d'origine*

Le premier moment de la représentation cartographique doit donc, en conséquence, être une analyse de l'information de départ pour déterminer ses composantes, et les hiérarchiser pour sélectionner celles qui seront préservées dans la simplification pour la représentation.

Pour simplifier utilement, il faut analyser l'information et fixer un objectif à la représentation.

En dehors des principes généraux qui guident la représentation (être clair et objectif, thématique et quantitatif), il faut fixer un objectif à la carte que l'on va réaliser, quant à l'information qu'elle doit *transmettre*.

Cet objectif servira de guide lors de la sélection des caractéristiques de l'information à garder lors de la simplification et surtout lors du choix du type précis de représentation à effectuer.

En fait, le principe d'objectivité reste le guide principal, mais le nombre de techniques de représentation disponibles laisse de la marge à l'expression de l'information.

Cette opération de regroupement en classes (qui sera détaillée dans la leçon sur les cartes en plages de couleurs) fait perdre énormément de l'information de départ, du point de vue

numérique. On passe, dans le cas d'une carte des départements français de 95 valeurs à 6 ou 7 seulement. C'est donc une simplification de l'ordre de 90% qui est effectuée !

Par exemple, lors de la réalisation d'une carte concernant une variable relative comme la densité de population, on est obligé d'utiliser une carte en plages de couleurs (cf. la leçon sur l'analyse de l'information), et l'on va donc simplifier l'information de départ en regroupant en classes les valeurs d'origine, ces classes devenant les valeurs, les couleurs, de la carte.

Comment, lors d'une telle simplification, préserver l'information importante ? Le sens de la variable numérique ?

Il existe plusieurs méthodes de découpage en classe (les *discrétisations*), qui seront examinées dans le détail plus loin, mais comment faire le choix entre tel ou tel découpage ?

➪ Il faut fixer un objectif à la représentation, et vérifier au long de la procédure que le résultat prévu reste conforme à cet objectif.

Pour reprendre l'exemple de la cartographie des densités dans les départements français, on peut fixer plusieurs objectifs à la carte, suivant l'utilité que cette dernière doit avoir (de manière caricaturale, bien sur) :

Auteur de la carte	Utilité de la carte	Objectif 1	Objectif 2
Chercheur Géographe	Etude, exploration	Avoir un aperçu global, déterminer des groupes spatiaux	Repérage des atypicités, des originalités
Gouvernement	Démonstration	Montrer l'homogénéité, l'harmonie	Désactiver les revendications potentielles
Elu local	Démonstration	Montrer l'hétérogénéité	Valoriser/dévaloriser son département

Ainsi, la simplification nécessaire pour une bonne compréhension de la représentation graphique amène parfois à un choix assez difficile, que seule une réflexion préalable sur l'utilité de la carte peut aider.

Nous nous placerons, dans ce cours, dans la position neutre, visant à minimiser la perte d'information, et privilégieront l'objectif de la réalisation de cartes d'aide à l'étude, d'exploration.

PRÉPARATION DE LA VARIABLE A REPRÉSENTER

Adaptation du niveau d'agrégation spatiale à l'objectif de la carte

Avant de se lancer dans la partie automatisée de la représentation cartographique, il faut fournir au logiciel que l'on utilise plusieurs types d'informations :

➪ *Un fond de carte numérique adapté à la représentation;*
➪ *Une ou plusieurs séries de données numériques contenant l'information à représenter.*

La question du fond de carte sera traitée dans une leçon suivante, nous nous intéresserons toutefois ici au niveau d'agrégation du fond de carte, parfois aussi appelé de manière assez ambiguë "échelle" du fond de carte. En effet, une fois l'information récoltée (cf. la liste de liens Internet en annexe), l'objectif de la carte fixée, il faut se poser la question suivante :

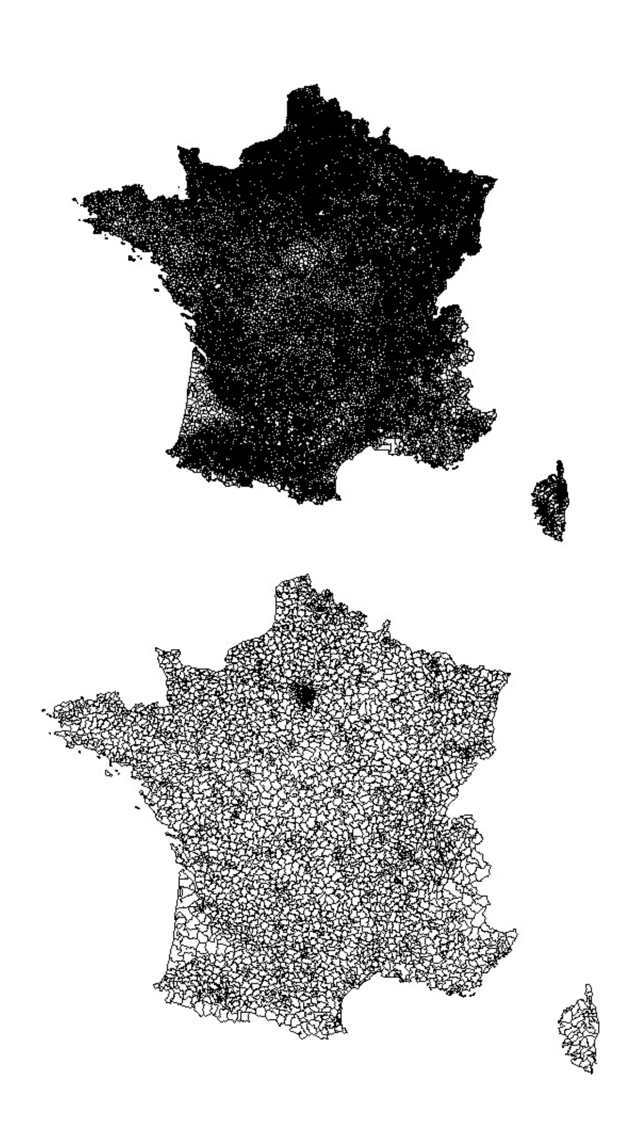
➪ Le niveau de précision géographique des informations récoltées est-il adapté à l'objectif fixé à la carte ?

Souvent (et c'est tout a fait sain) la récolte de l'information numérique se fait au niveau le plus fin possible, au niveau auquel l'information à été produite par l'organisme de statistiques ou l'enquêteur sur le terrain.

Or, on l'a vu, la représentation cartographique demande une simplification de l'information pour la rendre plus facilement saisissable, et il faut donc se poser la question de la simplification du niveau d'agrégation de l'information de base pour faciliter sa lecture.

Par exemple, pour reprendre l'idée de la carte des densités en France, l'information est produite par l'INSEE au niveau communal, voire au niveau des quartiers dans les agglomérations. Mais une carte dans un format courant ne peut présenter clairement les 36.600 communes de France, et ce niveau d'information, cette précision, n'est pas compatible avec un objectif de repérage rapide de la répartition des densités de population sur le territoire national.

Il faut donc envisager de regrouper les informations à un niveau d'agrégation spatiale supérieure : les cantons, les arrondissements, les départements...





Cette opération se fait généralement soit en allant chercher l'information directement au niveau d'agrégation voulu, soit en agrégeant l'information fine avec des outils comme MS Excel© ou un logiciel gestionnaire de Bases de Données.

Adaptation du niveau d'agrégation au fond de carte

La récupération de variables statistiques et leur transformation sont des opérations qui sont devenues progressivement moins difficiles avec l'apparition de la micro-informatique. Les organismes de statistiques diffusants leurs données sous forme numériques, voire sur leurs serveurs Internet, et les outils de traitement numérique et statistiques sont de plus en plus courants et accessibles.

Par contre, la disponibilité d'un fond de carte numérique adapté précisément aux données numériques disponibles n'est pas encore très courante.

En effet, la création d'un fond de carte numérique est encore une opération assez complexe, qui revient souvent à dessiner à la souris le fond de carte nécessaire grâce à un logiciel spécifique. Alors que les formats de fichiers de données sont assez standardisés (txt, dbf…), les formats de fonds de cartes sont souvent liés à des logiciels propriétaires.

Ainsi, les logiciels de cartographie automatique sont diffusés avec un ensemble de fonds de cartes considérés comme les plus courants : France par département, par Régions, Monde par

Pays, etc. Parfois, la création d'un fond de carte adapté à ces logiciels demande l'utilisation de logiciels spécifiques, voire l'achat pur et simple.

Donc, avant de vous lancer dans la récolte d'informations tous azimuts sur un thème, veillez à recenser les fonds de carte qui vous seront accessibles ou à vérifier si la création de nouveaux fonds est envisageable.

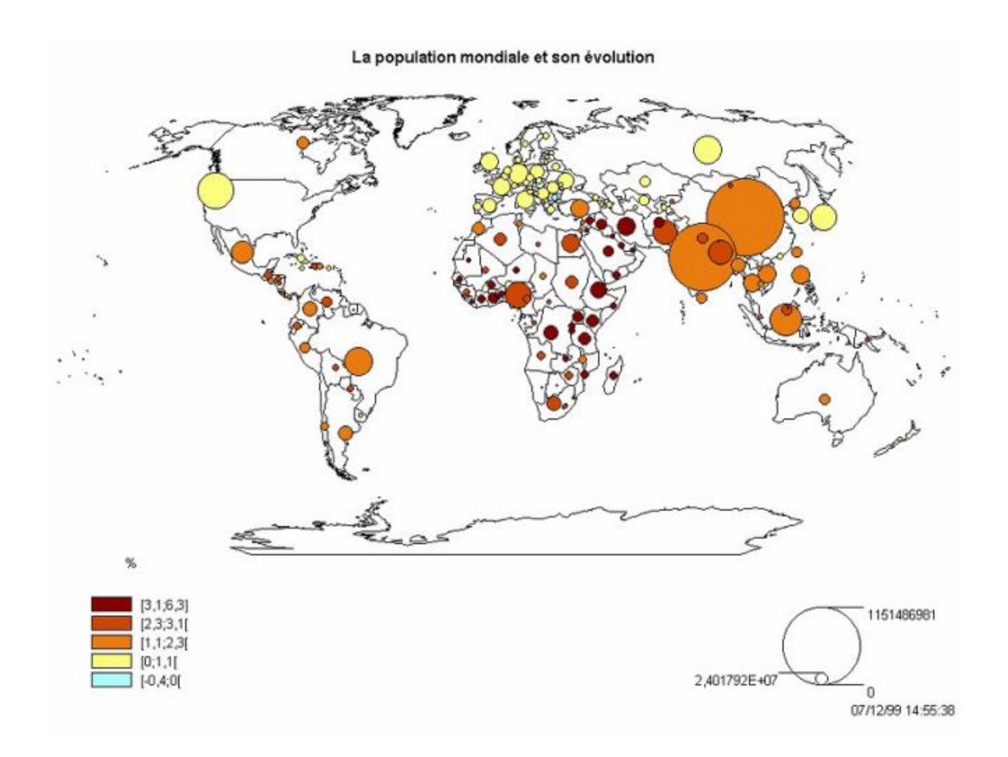
Sinon, vous serez obligés de modifier vos données statistiques pour les adapter au fond de carte disponible.

Adaptation de l'unité de la variable à l'objectif de la carte.

Lorsque vous cherchez à être synthétique sur une zone d'étude, essayez de réaliser des cartes qui elles aussi seront synthétiques et apporteront une information croisée. Les cartes et graphiques peuvent très bien se comparer aux paragraphes d'une argumentation : elles peuvent se compléter et se faire référence l'un l'autre.

Par exemple, au lieu de multiplier les cartes en valeurs absolues sur toutes les variables dont vous disposez (population, répartition par âges, par sexe, par activités... sur plusieurs années), essayez de croiser ces variables pour produire moins de cartes mais des cartes plus parlantes (densités, taux et rapports), qui viendront compléter une ou deux cartes en valeurs absolues. L'aboutissement de cette pratique serait alors les cartes de typologie, chaque type représentant une série de valeurs sur plusieurs variables.

On peut aussi envisager de réaliser une carte statistique double, présentant une variable absolue et une variable relative en même temps (la leçon suivante présentera les modalités du choix du type de carte à réaliser selon la variable, et une leçon est consacrée aux types de cartes complexes).



Il faut donc se poser la question suivante avant de se lancer dans la partie cartographique proprement dite :

Quelle est l'unité, la façon de présenter mathématiquement la variable qui sera la plus intéressante, par rapport aux autres cartes du document en cours de réalisation ?

Par exemple, est-il intéressant de représenter la population en 1990 et la population en 1999 ou la population en 1990 et le taux de variation entre ces deux recensements ?

Une fois l'objectif de la carte fixé, la variable adaptée, il faut se poser la question du type de carte à réaliser.

ANALYSE DE L'INFORMATION A REPRÉSENTER

Méthodologie du choix.

La Conception Cartographique Assistée par Ordinateur est un ensemble de techniques qui aident le géographe à représenter cartographiquement des données statistiques.

Souvenez-vous des principes énoncés en introduction :

> La carte à réaliser doit être la représentation la plus claire et la plus fidèle (objective) de l'information donnée.
> La carte à réaliser sera thématique et non synthétique.
> La variable à représenter est quantitative, donc ordonnée et proportionnelle.

A partir de ces principes et selon le type de variable spécifié, le cours de cartographie générale nous a appris que seuls certains types de cartes pouvaient être réalisés.

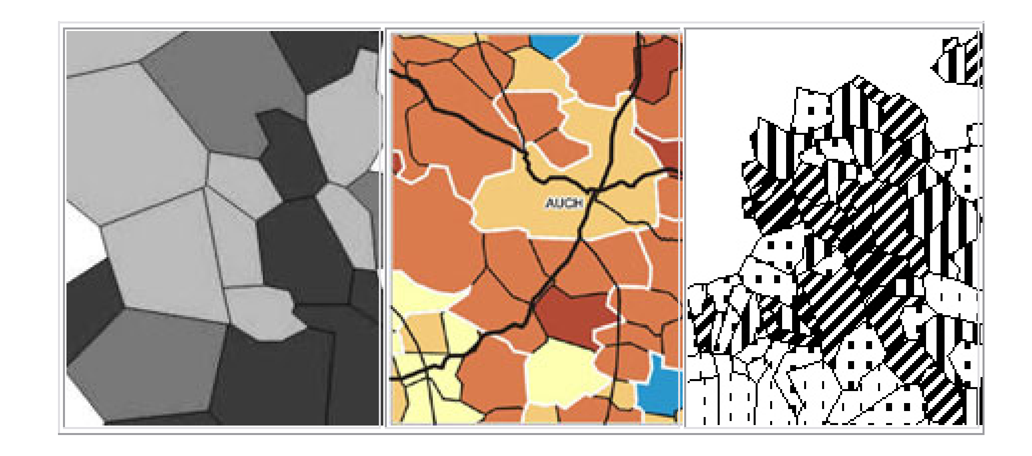
La première opération consiste à repérer quel est le type *d'objet cartographique* qui est la source d'information de la variable numérique à représenter. Ce repérage permet de choisir le type de fond de carte : points, lignes, ou polygones.

La seconde opération concerne le choix du type de représentation : *plages de couleurs* (carte dite Choroplèthe) ou carte en *Symboles Proportionnels* ?

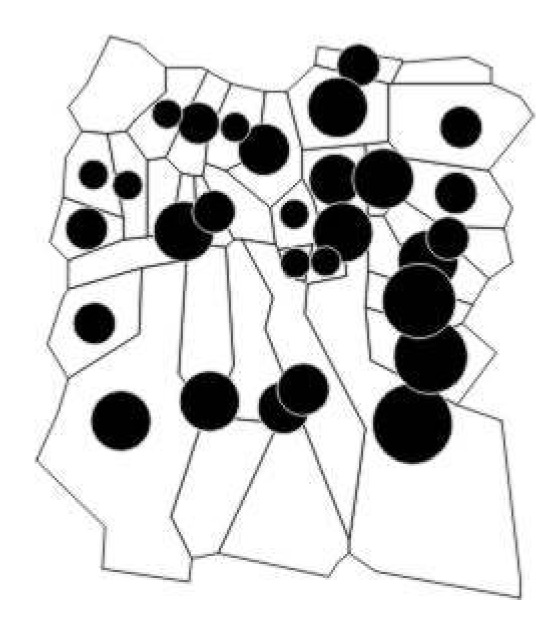
Le choix est plus ardu lorsque la source géographique de l'information est un objet que l'on peut représenter comme un point, une ligne, ou comme un polygone, une surface, et c'est très souvent le cas (unités administratives). La méthodologie de Jacques Bertin (sémiologie graphique) nous laisse le choix d'utiliser différentes techniques de dessin pour représenter l'information.

En effet, on peut représenter une variable statistique concernant une entité pouvant posséder une surface de deux manières :

- par une couleur, une valeur, une trame sur la surface occupée par cette entité;



- par un symbole ponctuel centré sur cette surface.



Il existe une règle, en CCAO, qui permet de sélectionner le type de carte à réaliser :

- Lorsque la variable statistique est absolue, brute, on utilise la variable graphique de la taille qui représentera directement ses variations quantitatives par les variations de la surface d'un symbole.
- Lorsque la variable statistique est relative, transformée, issue d'un calcul, elle est valable en tout point de la surface de l'objet géographique, on ne peut la relier à un objet concret. On utilise donc la variable graphique de la valeur et/ou de la couleur, adaptées aux informations différentielles ordonnées.

Le point clé de ce choix est donc une interrogation sur le type de variable statistique à représenter : absolue ou relative ? Brute ou transformée ? *La réponse à cette question est cruciale, car elle va déterminer le type de carte à réaliser.*

Synthétiquement :

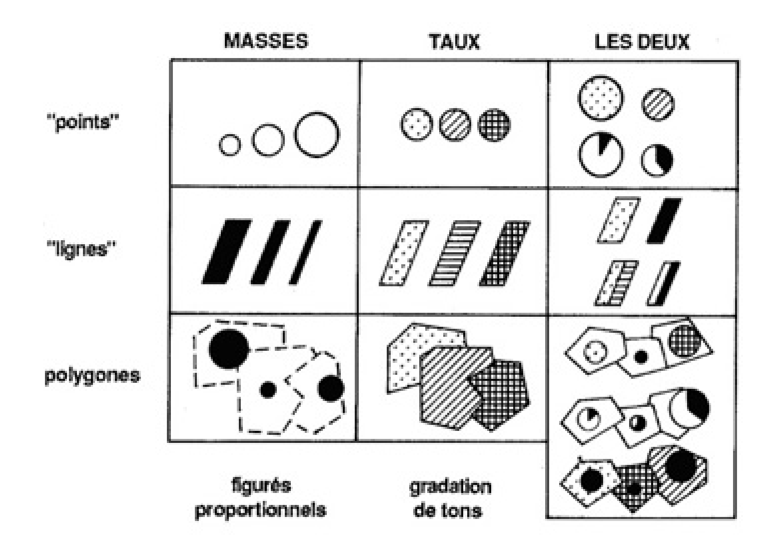
Type de Variable	Type de carte
Absolue, brute	Symboles proportionnels (taille)
Relative, transformée	Plages de couleurs (valeur/couleur)

Exemples de variables :

Absolue, brute	Relative, transformée
Population	Densité moyenne
Production	Taux d'évolution
Comptage de véhicules	Part de salariés dans la population active.

Les comptages, dénombrements, quantités brutes : habitants, tonnes, m², véhicules... sont des valeurs brutes, absolues. On représentera toujours, par exemple, les chiffres de la population sous la forme de symboles proportionnels.

Les moyennes, taux, indices, pourcentages sont des valeurs relatives (elles se définissent toujours en rapport a une autre valeur, ne sont pas autonomes). On représentera toujours les densités, par exemple, en plages de couleurs car elles s'expriment en habitants au km².



Masses et taux : les solutions disponibles

(Cf. Roger BRUNET, La carte mode d'emploi)

RÉALISATION DE CARTES EN SYMBOLES PROPORTIONNELS

Repérage des contraintes

Positionnement

Cette leçon est la première du chapitre concernant les cartes en symboles proportionnels.

Ce type de carte ne se réalise que dans la situation suivante :

- La variable statistique à représenter est absolue, brute, discrète.
- L'implantation des individus géographique sur le fond de carte est ponctuelle, ou rendue ponctuelle par la sélection du centroïde d'une implantation zonale.

Méthodologie

Une fois la variable et le fond de carte prêts, le type de carte à réaliser défini, il ne reste plus qu'à passer à la réalisation concrète de la représentation cartographique.

Celle-ci commence toujours par une recherche des différentes contraintes qui vont s'exprimer.

En effet, ce sont ces contraintes, qui, une fois repérées, vont guider les choix précis de représentation de la variable.

Les contraintes de la variable

Le premier groupe de contraintes provient de la variable même à représenter, et concerne principalement son étendue et sa dispersion :

- Repérage de l'étendue de la variable : différence entre la valeur la plus grande et la valeur la plus faible.

Cette étendue est une contrainte, car pour représenter la variable en utilisant la surface d'un symbole, il faut choisir un *ratio valeur de la variable / surface du symbole* qui permette de représenter correctement tous les individus.

Concrètement, il faut pouvoir repérer facilement sur la carte l'individu le plus faible et l'individu le plus important, même si ils sont très éloignés par leur valeur. Nous verrons dans la leçon suivante quelles techniques permettent de s'en sortir dans ce cas-là.

Si cette contrainte n'est pas correctement repérée, on risque de se retrouver à faire une carte où les plus petits symboles n'apparaîtront pas !

 Repérage de la dispersion de la variable : répartition des valeurs.

La dispersion des valeurs à l'intérieur d'une variable peut jouer un rôle important sur l'aspect général de la carte en symboles.

Si la dispersion des valeurs est faible, on risque de se retrouver avec une carte où les symboles ont à peu près tous la même taille, mis à part les individus extrêmes.

Au contraire, avec une dispersion trop grande, il risque de n'y avoir que des petits et des gros symboles sur la carte.

Dans ces cas extrêmes, ces problèmes de dispersion doivent être pris en compte et le ratio valeur / surface doit être transformé.

Les contraintes du fond de carte

Même si la variable semble à peu près équilibrée dans son étendue et sa dispersion, on ne peut imaginer à l'avance l'aspect général de la carte que si on s'intéresse de plus près à l'autre élément principal : le fond de carte.

En effet, les individus géographiques, comme les individus de la variable, peuvent présenter des contraintes importantes, dans leur répartition dans l'espace.

La principale contrainte concernant le fond de carte pour les analyses symboliques concerne en effet la répartition des points qui vont supporter des symboles dans le plan.

 Repérage d'éventuelles contraintes de répartition

Les contraintes de répartition sont les contraintes issues de la grande proximité d'individus à valeurs élevée, ce qui risque de nuire à la lisibilité de la carte.

Elles souvent liées à la présence de grandes villes dans le fond de carte. Souvent, les villes centres des agglomérations possèdent des valeurs très importantes, de loin supérieures à celles des communes qui les entourent. Le problème va être alors de représenter sur la même carte ces valeurs très différentes et très rapprochées.

A d'autres échelles, cette contrainte apparaît aussi, dans les cas de fonds de cartes possédant une forme allongée, comme dans le cas du Viêt-nam ou du Chili, par exemple.

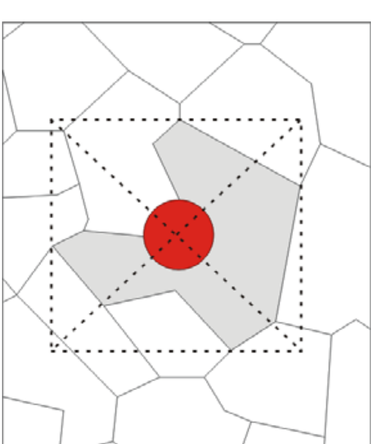
Repérage de problèmes de positionnement des centroïdes

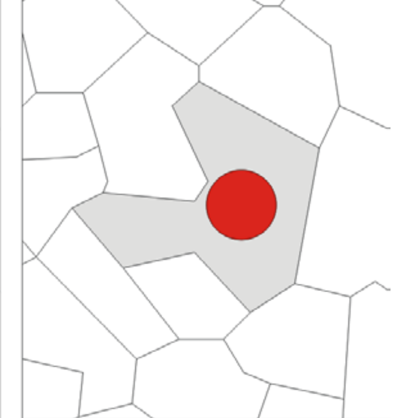
L'autre grande contrainte du fond de carte pour les cartes en symboles proportionnels provient de la nécessité de localiser correctement un point à l'intérieur d'une surface.

En effet, les fonds de carte informatisés sont souvent des fonds composés de polygones, de surfaces, et le logiciel va représenter les analyses symboliques en plaçant les symboles sur un point à l'intérieur des polygones, nommé centroïde.

Selon les logiciels, il est possible d'utiliser des centroïdes automatiques ou d'utiliser un fond de carte de points en superposition avec celui de polygones, pour utiliser par exemple la localisation exacte des villes préfectures pour les départements.

Mais le plus souvent le logiciel détermine seul la position de ce point, avec plus ou moins de bonheur, car souvent c'est la méthode du barycentre qui est utilisée.





Les contraintes de réalisation et de diffusion

Les contraintes de réalisation sont celles qui vont limiter les possibilités de représentation à cause de manques ou de défaillances des outils utilisés, ainsi que les contraintes qui apparaissent lorsque l'on réalise une série de cartes.

Ces contraintes doivent si possible être connues à l'avance pour éviter de se lancer dans des travaux qui ne pourront être concrétisés par une belle carte...

Par exemple, certains logiciels ne savent pas calculer correctement un ratio surface/symbole et préfèrent utiliser des symboles dits "cliparts" à la place de symboles mathématiques simples. Ce cas est très répandu et des logiciels aussi courants que Cartes & Bases© ou Mapinfo© sont concernés. La volonté de mettre à la disposition de l'utilisateur une vaste bibliothèque de symboles différents a primé sur la rigueur méthodologique de la représentation.

Les logiciels de cartographie automatiques présentent un ensemble de possibilités variable d'un logiciel à l'autre, selon l'orientation du logiciel (marketing, comme MsMap Point©) ou scientifique (comme Cabral© ou SCAP©).

Ainsi, avant de réaliser les travaux préparatoires de la carte, il faut essayer de repérer si le logiciel dont on dispose est capable de prendre en compte tous les besoins relevés lors des étapes précédentes : détermination de l'objectif, analyse de la variable, repérage des contraintes.

Les contraintes de série sont liées au fait que lorsque l'on réalise un document devant contenir plusieurs cartes, il faut suivre certaines règles pour essayer de produire un document harmonieux dans sa présentation :

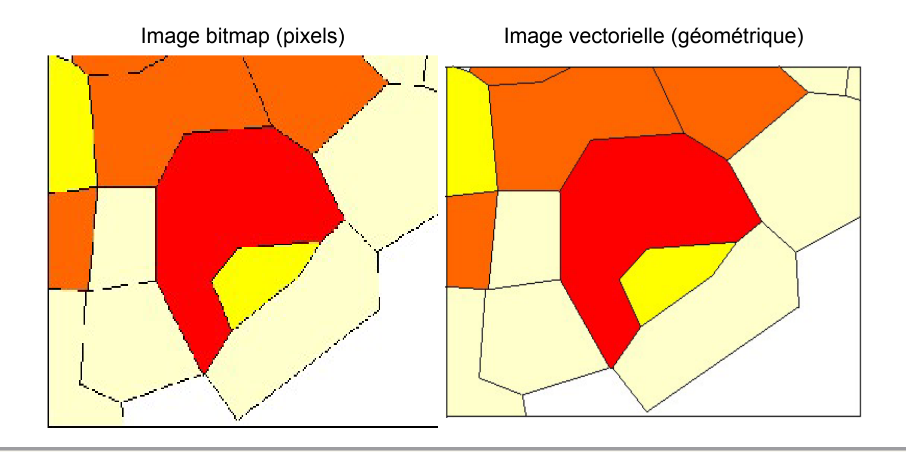
- Utiliser toujours les mêmes styles de mise en page et de présentation pour les cartes
- Utiliser des couleurs différentes pour les symboles et si possibles liées à la variable étudiée (eau=bleu, forêt=vert, etc.)
- Dans le cas de cartes en vis à vis et/ou complémentaires, utiliser si possible le même ratio valeur/surface, pour que les légendes soient comparables directement.

Les contraintes de diffusion regroupent les limites à la créativité induites par le type de technique d'impression et de diffusion (reproduction) de la carte.

En effet, c'est trivial, mais on voit encore trop souvent de belles cartes couleurs perdre toute leur efficacité et tout leur intérêt par une mauvaise photocopie en noire et blanc très contrastée.

Autre problème possible, certains logiciels savent bien faire les cartes mais ont du mal à les exporter convenablement, dans des formats compatibles avec d'autres logiciels, et notamment les logiciels de mise en page et de publication assistée par ordinateur.

Les cartes à l'écran du logiciel sont en effet sous la forme de points colorés de l'écran, les pixels. Ces points ne peuvent pas être utilisés pour un agrandissement de la carte pour l'impression par exemple, car on voit rapidement apparaître la forme des points qui composent l'image :



Synthèse méthodologique

Mais la contrainte principale lors de la réalisation d'une carte en symboles proportionnelle est issue à la fois de l'étendue de la variable et de la forme du fond de carte : c'est la taille du plus gros symbole que l'on peut dessiner sur le fond de carte, sans que le symbole ne masque le fond. Cette taille dépend naturellement de l'étendue de la variable et de la forme du fond de carte, ainsi que de la répartition des valeurs à l'intérieur de ce fond, a priori.

Ensuite, viennent s'appliquer les autres contraintes :

- Taille du plus gros symbole possible suivant la forme du fond et les contraintes d'agglomération
- Conséquence : Taille du plus petit symbole visible sans effort

- Dispersion des valeurs dans la variable
- Contraintes techniques de réalisation et de diffusion de la carte.

Une fois ces contraintes repérées et notées, on peut passer à la phase de cartographie proprement dite.

RÉALISATION DE CARTES EN SYMBOLES PROPORTIONNELS

Paramétrage et Cartographie

Conseils généraux de réalisation

Cette série de guides est issue de l'expérience synthétisée dans les excellents ouvrages de Edward R. Tufte (cf. la bibliographie).

- La carte doit avant tout montrer la variable, c'est l'information qui doit être perçue en premier lorsqu'on lit la carte.
- La carte doit avoir un objectif clair : description, exploration, mise en forme, illustration.
- La carte doit inciter le lecteur à se poser des questions à propos de la répartition spatiale de la variable, pas à propos de la méthode/technique de représentation utilisée.
- Essayez de réduire au maximum la distorsion de la variable induite par la représentation graphique.
- Essayez d'utiliser le moins de place possible pour présenter le plus d'information.
- La carte doit montrer la cohérence de la variable présentée.
- Essayez de montrer la variable à différents niveaux d'interprétation : globale, locale, structure fine.
- La carte doit être intégrée au document, liée aux textes qui l'entourent, servir une argumentation, une démonstration.

Technique du ratio valeur / surface : 1° - La taille du plus gros symbole.

Comment passer concrètement d'une série de valeurs numériques à un ensemble de symboles dessinés sur un fond de carte ?

Cette opération s'effectue automatiquement dans le logiciel de CCAO (et c'est là toute l'utilité de l'automatisation, avec le dessin) par le calcul d'un ratio, d'un coefficient, qui, une fois appliqué à une valeur de la série donne la *surface* du symbole la représentant graphiquement.

Attention, ici c'est bien la surface du symbole qui doit être proportionnelle à la valeur, car ce que l'on perçoit lorsqu'on regarde une carte c'est la présence de "taches " plus ou moins grandes qui se découpent sur le fond de la carte, et pas une mesure plus abstraite comme le rayon ou le diamètre des symboles.

Pourquoi ce rappel de techniques apprises lors des leçons du cours de Cartographie manuelle ? Tout simplement parce que les logiciels de CCAO ne savent pas forcément calculer des proportionnalités à la surface et réaliser a priori quelle sera la taille adéquate des symboles, donc la valeur du ratio.

En effet, les logiciels de CCAO attendent que vous leur fournissiez une correspondance entre une valeur de la variable et une taille de symbole (une surface de symbole). A partir de cette correspondance, par exemple 12,2 -> 200 pixels², ils vont en déduire le ratio et l'appliquer à toutes les autres valeurs. C'est beaucoup plus simple (parce que cela n'implique pas de calculs, juste un choix) que de demander à l'utilisateur un ratio.

➡ C'est donc à vous de choisir la valeur de référence et la surface de symbole que vous allez lui affecter.

Comment choisir ces valeurs ? Tout simplement en tenant compte des contraintes repérées précédemment.

En effet, c'est à ce moment de la méthodologie que les contraintes identifiées vont nous être utiles. Quelle est la principale contrainte de la représentation en symboles proportionnels : *la taille maximale de symbole suivant le fond de carte*.

➡ Le choix de la valeur de référence du ratio valeur/surface doit donc se porter sur la valeur maximale de la variable, puisque l'on peut définir a priori quelle est la surface maximale du symbole.

Ainsi, on recherche où va se trouver la valeur maximale, et on évalue la taille maximale du symbole que l'on peut dessiner à cet endroit sans trop cacher le fond de carte.

Exemple :

Dans le cas d'un travail "à la main", on aurait procédé de cette manière, pour le dessin de cercles proportionnels :

Valeur	Surface	Rayon
987	200	7,98
457		
213		
124		
56		
10		

$$S = \pi . r^2$$

$$r = \sqrt{\frac{S}{\pi}}$$

On choisit la valeur de 200 pixels² pour la taille maximale de symbole, après avoir repéré que le rayon maximal dessinable était de 8mm.

*On utilise ensuite ce ratio 987->200 pour calculer les surfaces des autres symboles (en proportionnalité simple, règle de trois : 457 * 200 / 987 = 92,6), puis, à partir des surfaces, on obtient les rayons qui vont nous permettre de dessiner les symboles.*

Valeur	Surface	Rayon
987	200	7,98
457	92,6	5,43
213	43,2	3,71
124	25,1	2,83
56	11,3	1,90
10	2,0	0,80

Dans le cas d'un travail en CCAO, il n'y a plus besoin de calculer à la main les valeurs de chaque rayon, ni de les dessiner. Mais il faut toujours définir quelle va être le ratio qui sera utilisé pour passer des valeurs de la variable à une taille de symbole sur la carte.

Une fois le paramètre de liaison valeur/surface entré dans le logiciel, la rapidité d'exécution du dessin de la carte peut être utilisée pour vérifier rapidement, sans préciser pour l'instant les autres paramètres de la représentation, si le choix effectué pour la taille du plus gros symbole est correct.

Cette vérification permettra aussi de vérifier la seconde contrainte, issue de la première, c'est à dire la taille du plus petit symbole découlant du choix du ratio effectué précédemment.

Technique du ratio valeur / surface : 2° - La taille du plus petit symbole

C'est la manifestation concrète de la contrainte de dispersion des valeurs de la variable. Si la variable est trop dispersée, la surface du plus petit symbole découlant du choix du ratio risque de ne pas être utilisable.

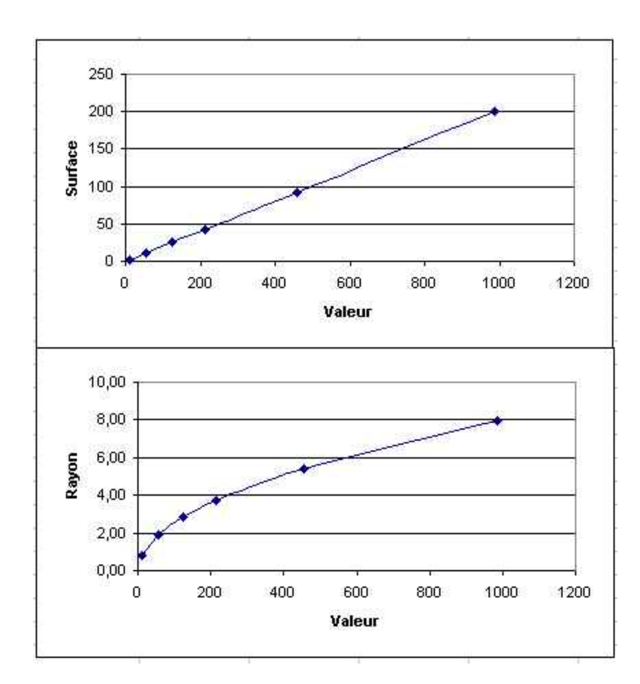
Deux solutions existent dans ce cas :

- l'utilisation d'un autre figuré pour le plus gros symbole, ce qui permet de le dessiner en utilisant une surface plus importante, donc de modifier le ratio (par exemple un figuré évidé);
- la transformation du ratio pour le faire passer d'un ratio simple, linéaire (valeur * ratio = surface), à un ration plus complexe, logarithmique ou racine (log (valeur) * ratio = surface ou racine (valeur) * ratio = surface).

Dans tous les cas, toute transformation du ratio vers un ratio non linéaire est à utiliser en dernier recours, car on perd alors la liaison simple entre la valeur et les surfaces du symbole, donc le lecteur de la carte, qui s'attend à une proportionnalité simple, risque d'être trompé. C'est pourquoi on l'indique toujours en légende lorsque c'est le cas.

Astuce : Si le logiciel de CCAO ne vous permet pas de transformer directement le ratio pour utiliser une fonction racine ou logarithmique, il suffit de ne pas utiliser les valeurs normales de la variable, mais leurs racine ou leur logarithme népérien. Les symboles seront alors transformés, mais la légende sera fausse, il faudra la corriger.

Lorsque le ratio est simple, linéaire, voila ce que donnent les courbes valeur/surface et valeur/rayon :



La première courbe est une droite : le ratio est linéaire. La seconde est une courbe qui ressemble à la courbe racine, c'est normal car il existe une proportionnalité carrée entre les surfaces et les rayons (S = Pi . R²).

Ratio Racine

Surface

250
200
150
100
50
0

0 200 400 600 800 1000

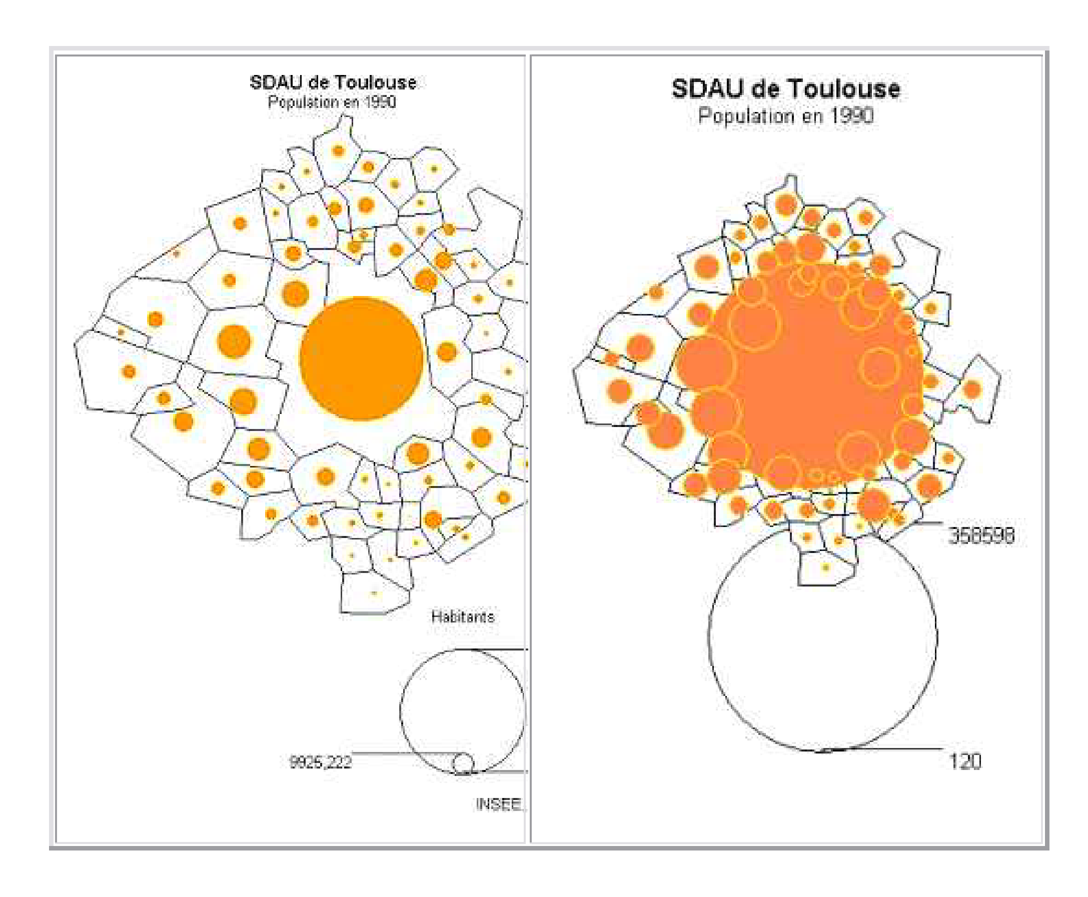
Valeur

Rayon

10,00
8,00
6,00
4,00
2,00
0,00

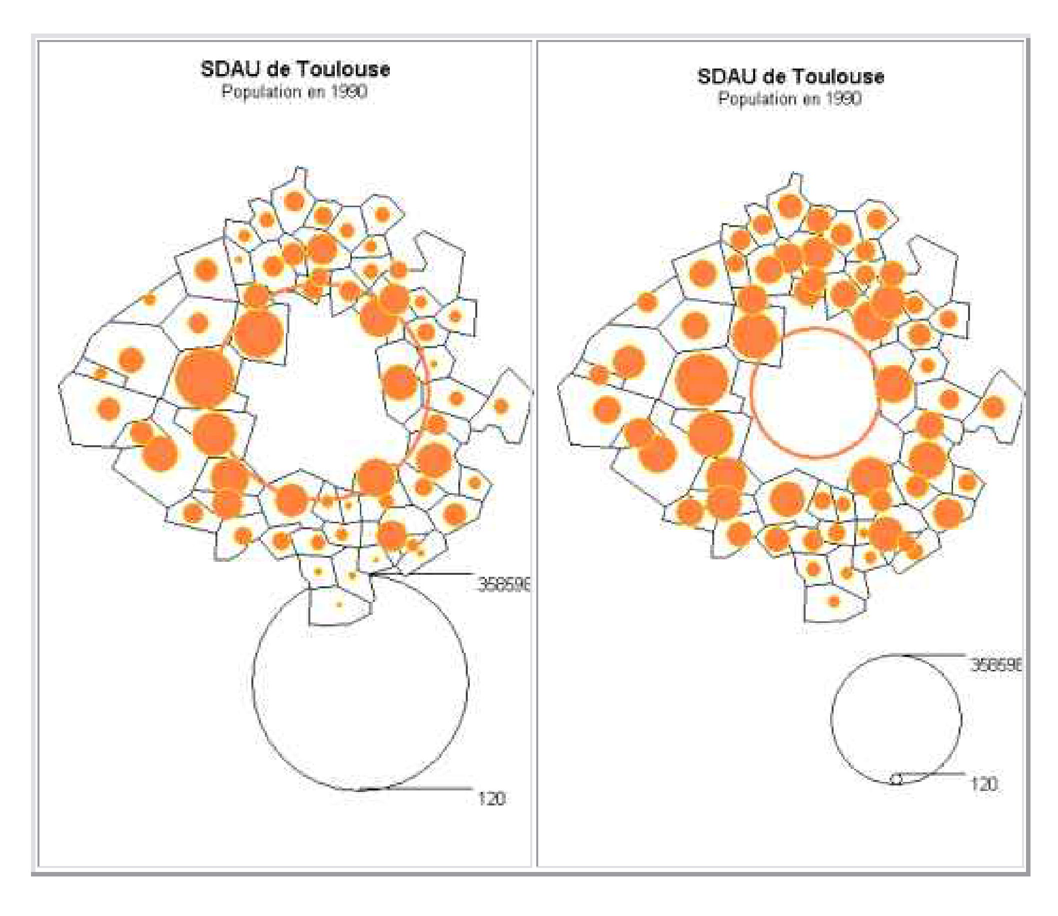
0 200 400 600 800 1000

Valeur



Lorsque le ratio est de type racine, la proportionnalité est transformée, la courbe valeur/surface n'est plus une ligne, mais une courbe racine, et la courbe valeur/rayon est un peu plus aplatie que dans l'exemple précédent. Lorsque le ratio est de type logarithmique, cette transformation est encore plus marquée.

Remarquez que le point (987, 8) reste toujours fixe sur les trois séries de graphiques : c'est l'origine du ratio, la plus grande valeur/le plus gros symbole. C'est la pente de la courbe qui mène à ce point qui varie.



Une fois le calibrage des symboles défini et vérifié, reste à choisir les paramètres qui vont fixer la présentation générale de la carte.

Paramétrage de l'aspect de la représentation des symboles

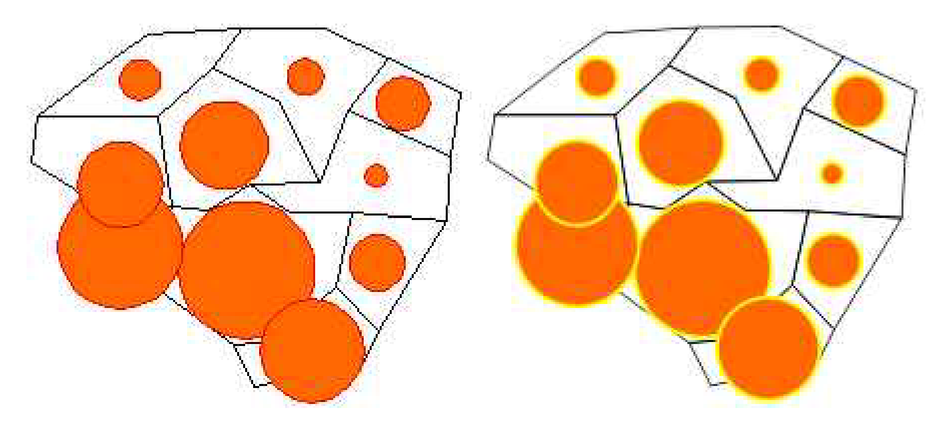
C'est à ce moment qu'interviennent les autres contraintes repérées lors de la leçon précédente : les contraintes de série et les contraintes de réalisation.

En effet, les logiciels de CCAO fonctionnant sur des écrans couleur, ils vous permettent d'utiliser les couleurs de votre choix et la finesse de trait la plus fine. Mais il va falloir tenir compte de la vie ultérieure de la carte :

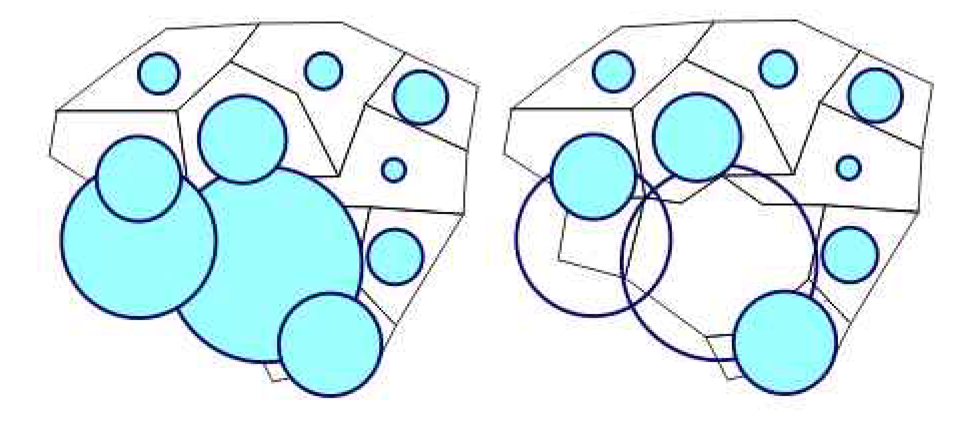
- Dans quel type de document cette carte va t'elle être incluse ?
  - taille maximale de la carte
  - nombre et type de couleurs utilisables
  - épaisseur minimale des lignes/contours.
- Fait-elle partie d'une série de documents graphiques ?
  - Quel camaïeu de couleurs est utilisé ?
  - Doit-on faire partager les légendes à plusieurs cartes ?

A partir de ces questions, on peut alors définir quelles couleurs utiliser pour les symboles et pour la matérialisation du fond de carte. Il existe certaines astuces de dessin pour présenter au mieux les symboles proportionnels :

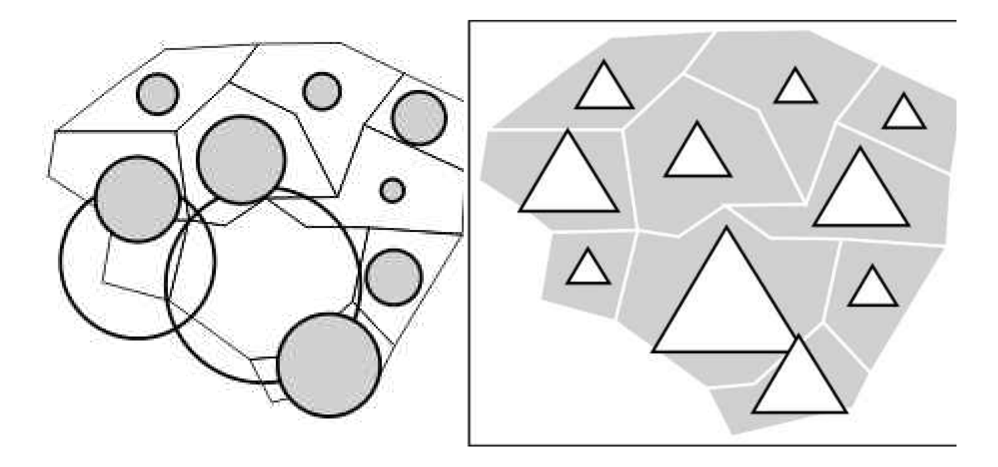
- Utiliser si possible des couleurs bien contrastées pour le contour et le fond des symboles (pour bien repérer les symboles qui se superposent).



- Dans le cas de superpositions trop importantes et/ou de symboles trop grands, essayez d'utiliser des symboles évidés.



- Dans le doute sur la possibilité d'utiliser la couleur, ne soyez pas timide d'utiliser le noir est blanc, qui, lorsque les nuances sont bien contrastées est très lisible.



RÉALISATION DE CARTES EN SYMBOLES PROPORTIONNELS

Présentation de la légende et mise en page

A quoi sert une légende ?

La légende est en fait la clé de lecture d'une carte.

Elle doit donner tous les renseignements nécessaires à la bonne compréhension de l'information de la carte, et ne pas induire en erreur le lecteur par des informations incomplètes ou ambiguës.

Une légende doit donc comporter :

- Un titre précisant l'unité dans laquelle est exprimée la variable.
- Une présentation graphique de la liaison variable / surface de symbole pour quelques valeurs représentatives de la variable, si possible comprenant le minimum et le maximum.

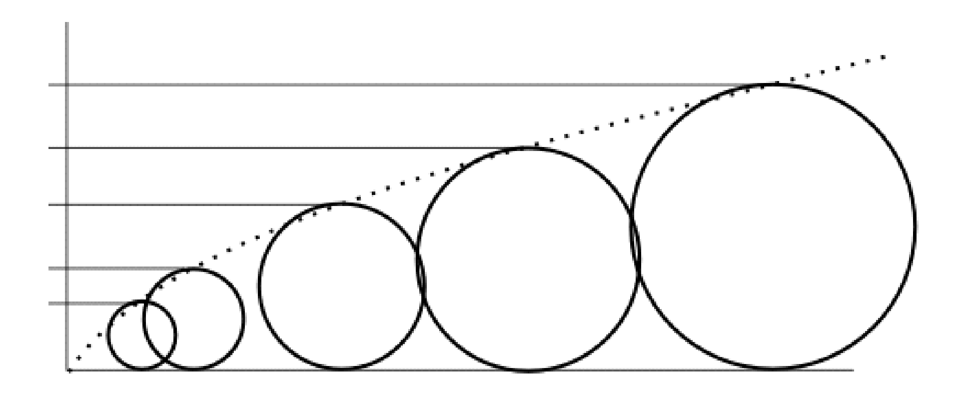
Ce n'est pas la peine de répéter le nom de la variable, le nom de la zone géographique représentée, la date de l'information en légende : c'est le rôle du titre de la carte.

En effet, une carte scientifique se doit d'être simple et claire, donnant toutes les clés de lecture dans le titre et dans la légende, sans commentaires superflus.

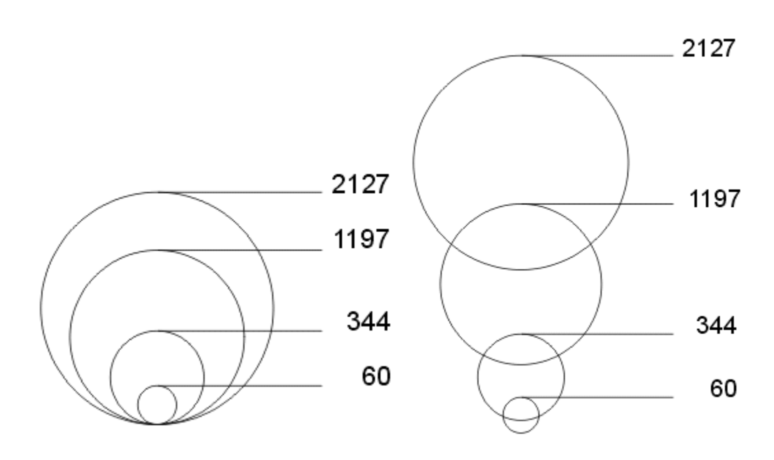
Présentation de la légende

Le problème des légendes des cartes en symboles proportionnels est qu'elles doivent présenter des valeurs représentatives de la variable. Souvent, cela se traduit par une légende prenant une place assez conséquente sur la page.

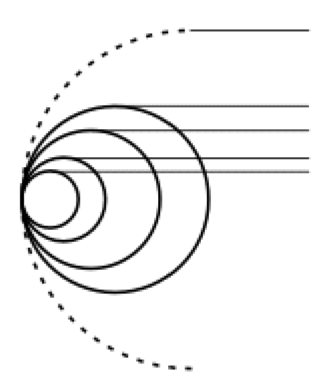
La présentation la plus claire est celle qui reprend l'idée d'un graphique, car elle matérialise la forme du ratio utilisé et informe donc le lecteur sur une éventuelle transformation du ratio :



Mais cette présentation a le défaut d'occuper beaucoup de place. On peut réduire la taille de la légende en évacuant le dessin de la courbe du ratio valeur/surface.



Parfois, le plus gros symbole est trop grand pour être répété en légende, on peut alors soit le présenter en demi-symbole avec une teinte plus claire pour ne pas encombrer la carte, soit écrire en toutes lettres sa valeur en petits caractères.



Mise en page générale

La mise en page d'une carte comprend certains éléments que l'on placera autour du fond de carte. Mais retenez qu'une carte claire et agréable à lire est une carte simple, sans fioritures.

Une carte doit présenter, autour du fond de carte :

- Un titre qui répond aux questions : Quoi, Où, Quand ?
- Une légende qui renseigne sur l'unité utilisée et le ratio variable/surface du symbole
- Une échelle graphique (cf. exemples ci-dessous)
- Une flèche indiquant l'orientation de la carte si celle-ci n'est pas évidente (nord)
- Une indication de la source de la carte (pour informer sur la validité des données et leur date de mise au point)
- Une indication de l'auteur de la carte si celui-ci n'est pas celui du document global ou si la carte ne fait pas partie d'une série.

**RÉALISATION DE CARTES
EN PLAGES DE COULEURS**

Recherche des contraintes

Positionnement

Cette leçon est la première du chapitre concernant les cartes en plages de couleurs, dont le nom scientifique est choroplèthe.

Ce type de carte ne se réalise que dans la situation suivante :

- La variable statistique à représenter est relative, continue.
- L'implantation des individus géographique sur le fond de carte est zonale, ou possède une composante possédant une surface (exemple des cartes mixtes, cf. la leçon sur les autres types de cartes).

Méthodologie

Une fois la variable et le fond de carte prêts, le type de carte à réaliser défini, il ne reste plus qu'à passer à la réalisation concrète de la représentation cartographique.

Celle-ci commence toujours par une recherche des différentes contraintes qui vont s'exprimer.

En effet, ce sont ces contraintes, qui, une fois repérées, vont guider les choix précis de représentation de la variable.

Une fois les contraintes identifiées, on pourra passer aux choix concernant les paramètres a priori et les paramètres spécifiques de la symbolique.

Les contraintes de la variable

Le cas des cartes en plages de couleurs est particulier, car on a affaire avec une variable quantitative continue à représenter sur une surface.

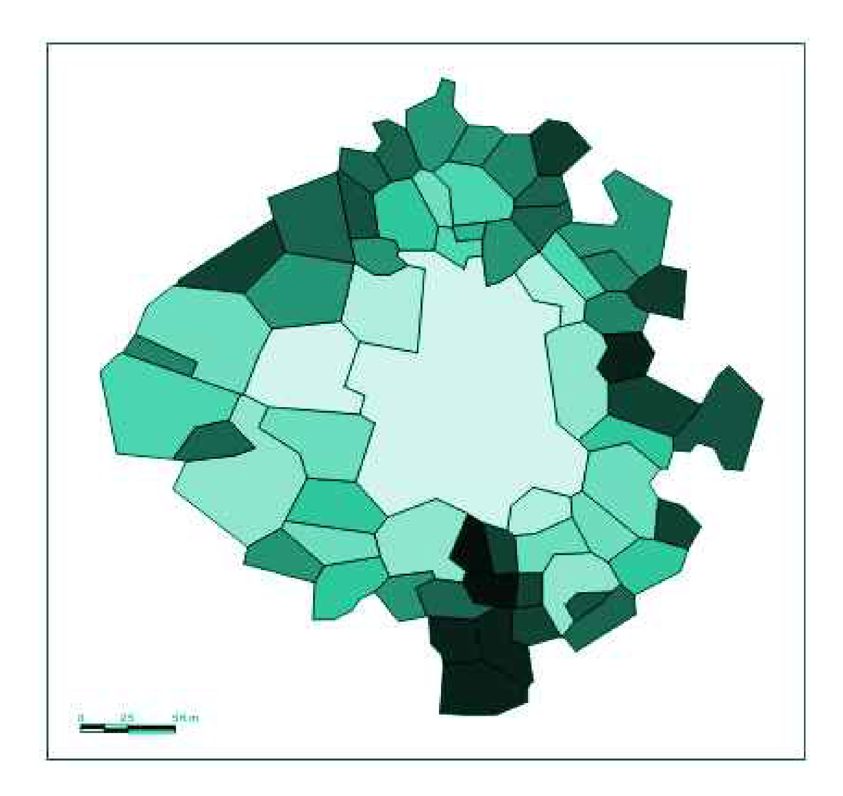
La sémiologie graphique de Jacques Bertin nous dit que les figurés zonaux à utiliser pour la meilleure performance sont les plages de valeur et par extension, de couleur. Ceci signifie que pour représenter au mieux une variable sur des surfaces, il est préférable d'utiliser des gradations de nuance d'une couleur principale.

La contrainte majeure qui se présente alors est celle de la vision humaine : en effet, si on appliquait le principe sémiologique directement, on obtiendrait alors une nuance différente par individu de la variable, c'est à dire autant que nuances de couleurs qu'il y a d'individus sur la carte.

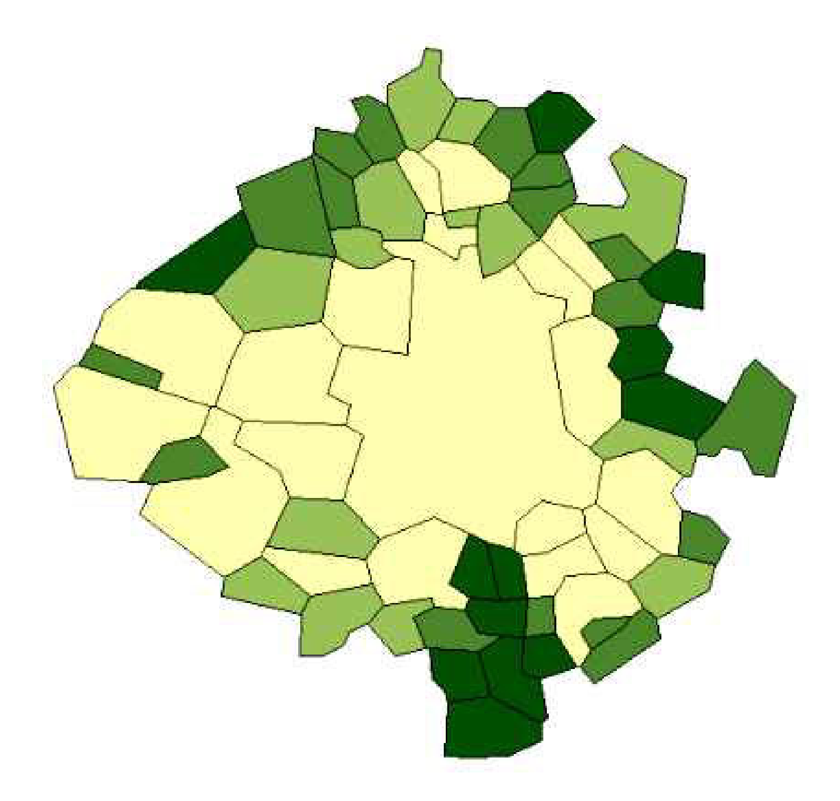
Or des chercheurs ont démontré que l'oeil humain ne distinguait (séparait) de façon globale qu'au maximum 7 nuances différentes d'une teinte.

Exemple :

Carte réalisée sans regroupement de valeurs



Carte réalisée en quatre classes :



La seconde carte, ci-dessus, ne possède que quatre couleurs différentes, et on retrouve plus facilement toutes les communes de la même nuance, appartenant à la même classe. L'analyse de la carte est donc plus rapide, plus simple.

L'objectif principal d'une carte n'est pas d'être agréable à regarder, mais de faire passer une information.

Comment transformer une variable possédant x individus en une variable d'au maximum 7 groupes d'individus ? : en réalisant une classification, une discrétisation.

En effet, l'opération consistant à rendre discrète (avec un nombre précis de valeurs possibles) une variable continue (qui peut prendre n'importe quelle valeur dans une intervalle donné) s'appelle la discrétisation, et fera l'objet de la leçon prochaine.

La variable elle-même peut ensuite poser d'autres contraintes. L'étendue et la dispersion ne sont pas ici des problèmes majeurs, ils seront à prendre en compte dans la création des classes.

Par contre, comme la transcription de l'information est basée sur des gammes de nuances, cette progression dans la couleur implique une progression dans les valeurs des classes de la variable. Or on n'a pas toujours à faire à des variables présentant une seule progression dans ses valeurs.

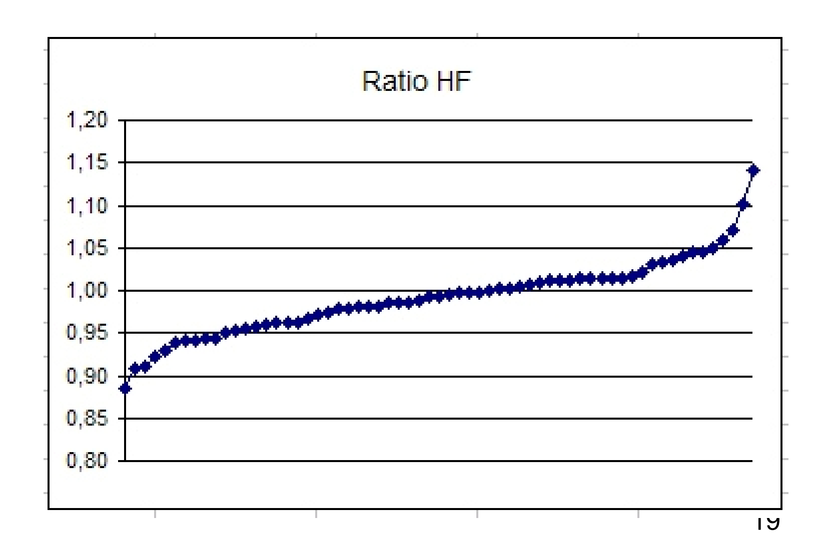
Il faut prendre en compte la présence éventuelle de valeurs négatives, et/ou de valeurs seuils.

En effet, le guide principal de la représentation graphique étant de retranscrire le plus fidèlement possible la variable, il va falloir s'assurer si cette dernière possède des flux internes, des gammes de valeurs importantes.

La plus classique de ces situations concerne la présence de valeurs négatives : on a alors deux progressions à partir de zéro, vers les valeur élevées et vers les valeurs faibles. N'oublions pas que les valeurs à représenter sont des valeurs relatives, des taux, on aura donc souvent affaire à des valeurs négatives.

Parfois, les variables possèdent une valeur seuil, une valeur de référence : l'indice 100, la moyenne, un taux plancher, etc. Ces valeurs seuils peuvent servir à découper la variable en deux sous-groupes, inférieur et supérieur. Ce qui rendra la lecture de la carte plus simple et surtout plus informative.

Par exemple, voici la courbe des ratios Hommes / Femmes triés, la valeur seuil est naturellement la valeur 1.



Les contraintes du fond de carte

En ce qui concerne les cartes choroplèthes, les contraintes imposées par le fond de carte sont moins importantes que pour les symboliques. En effet, le figuré étant ici directement dessiné à l'intérieur des polygones du fond, il n'y a pas d'effets de dépassement et de superposition à craindre.

Par contre, reste le problème des plus petits symboles qui doivent toujours être visibles. Or les plus petites surfaces sont souvent celles qui vont recevoir les teintes les plus claires (et les plus grands noms, c'est très courant)

Il faut que la teinte la plus claire de la gamme utilisée en légende soit nettement discernable de la couleur du fond de carte et facilement repérable sur la carte, même si elle se trouve dans de petites surfaces.

Ainsi, ne sélectionnez pas un découpage en classe aboutissant à une classe des valeurs les plus faibles ne possédant qu'un ou deux individus de très petite surface sur la carte.

Par extension, toujours pour la lisibilité, essayez de ne pas produire de cartes dont la répartition d'une classe n'est pas rapidement lisible (classes trop petite ou possédant des individus de trop faibles surfaces).

Les contraintes de réalisation et de diffusion

Comme la matérialisation de l'information statistique passe par une gamme de couleurs, il faut vérifier que la ou les teintes utilisées ne le soient que pour cet objectif. Séparer l'utilité de chaque couleur pour des objectifs différents permet de ne pas embrouiller la carte : une couleur neutre pour le fond de carte et l'habillage (légende, échelle, mentions), une couleur bien distincte et qui ressort bien pour l'information.

De par l'utilisation de figurés surfaciques, les contraintes de réalisation vont être importantes pour les choroplèthes. La sémiologie graphique nous indique que l'utilisation de la couleur est le meilleur moyen de retranscrire des variables différentielles, ce qui implique que l'ajout de différences de teinte à une gradation de nuance, de valeur, est intéressant.

Or la réalisation matérielle de cartes en couleur est encore assez problématique du point de vue technique, lorsque la diffusion du document se fait sur papier.

Le problème est que les logiciels de CCAO travaillant sur des écrans couleurs pouvant afficher des millions de nuances, et selon le principe des couleurs additives, ce que vous voyez à l'écran à de très fortes chances de ne pas ressembler beaucoup à ce qui va sortir de votre imprimante, qui utilise un système de couleurs très différent.

Non seulement la technique physique de représentation des couleurs change du tout au tout (couleurs soustractives), mais de plus la calibration des couleurs est un ensemble de techniques complexes difficiles à mettre en oeuvre sur du matériel grand public, même si ce dernier s'améliore progressivement.

Ainsi, votre écran peut être réglé d'une certaine manière, vos yeux percevoir les couleurs d'une manière différente et votre imprimante équilibrer les couleur d'une troisième façon...

Il faut donc encore une fois vous poser la question de l'avenir de votre carte : comment doit-elle être diffusée, reproduite, imprimée.

Nous verrons les astuces de dessin utilisables pour tenir compte de ces contraintes dans la leçon sur le paramétrage et la cartographie.

RÉALISATION DE CARTES EN PLAGES DE COULEURS

Le découpage en classes : la discrétisation

Position du problème

L'objectif d'une discrétisation est de ramener à au maximum sept classes (voire huit si l'on a deux progressions de couleur) un variable dont le nombre d'individus est changeant.

Les principes qui doivent guider la discrétisation visent à poursuivre les principes de la représentation graphique d'une variable statistique, le respect de l'information de départ en premier lieu :

- Respect de l'information de départ : la répartition des couleurs sur la carte, donc des individus dans les classes, doit être fortement liée à la répartition des individus dans la variable.
- Il ne doit pas y avoir de classe vide
- Il ne doit pas y avoir de classes non jointives (trou de valeurs entre deux classes successives)
- Les classes ne doivent pas se chevaucher (un individu ne peut appartenir qu'à une classe à la fois)

Il faut donc arriver à couvrir complètement la variable avec des classes jointives, tout en essayant de préserver la "forme" de l'information de départ.

Comment remplir ce premier objectif ? Comment savoir la forme de la répartition des individus dans la variable ? La méthode classique de la cartographie nous enjoint à observer de plus près la courbe des fréquences de la variable.

Le diagramme des fréquences

En CCAO, cette courbe des fréquences est réalisable de façon automatique, selon les logiciels (Cabral© et SCAP©), ce qui va nous permettre de mieux visualiser la forme d'une variable. Ainsi, on aura un point de comparaison avec les discrétisation que vont nous fournir les méthodes de classification automatiques.

Si vous utilisez un logiciel de CCAO ne possédant pas cette fonction de dessin du diagramme des fréquences, il est envisageable de le réaliser sous Excel©.

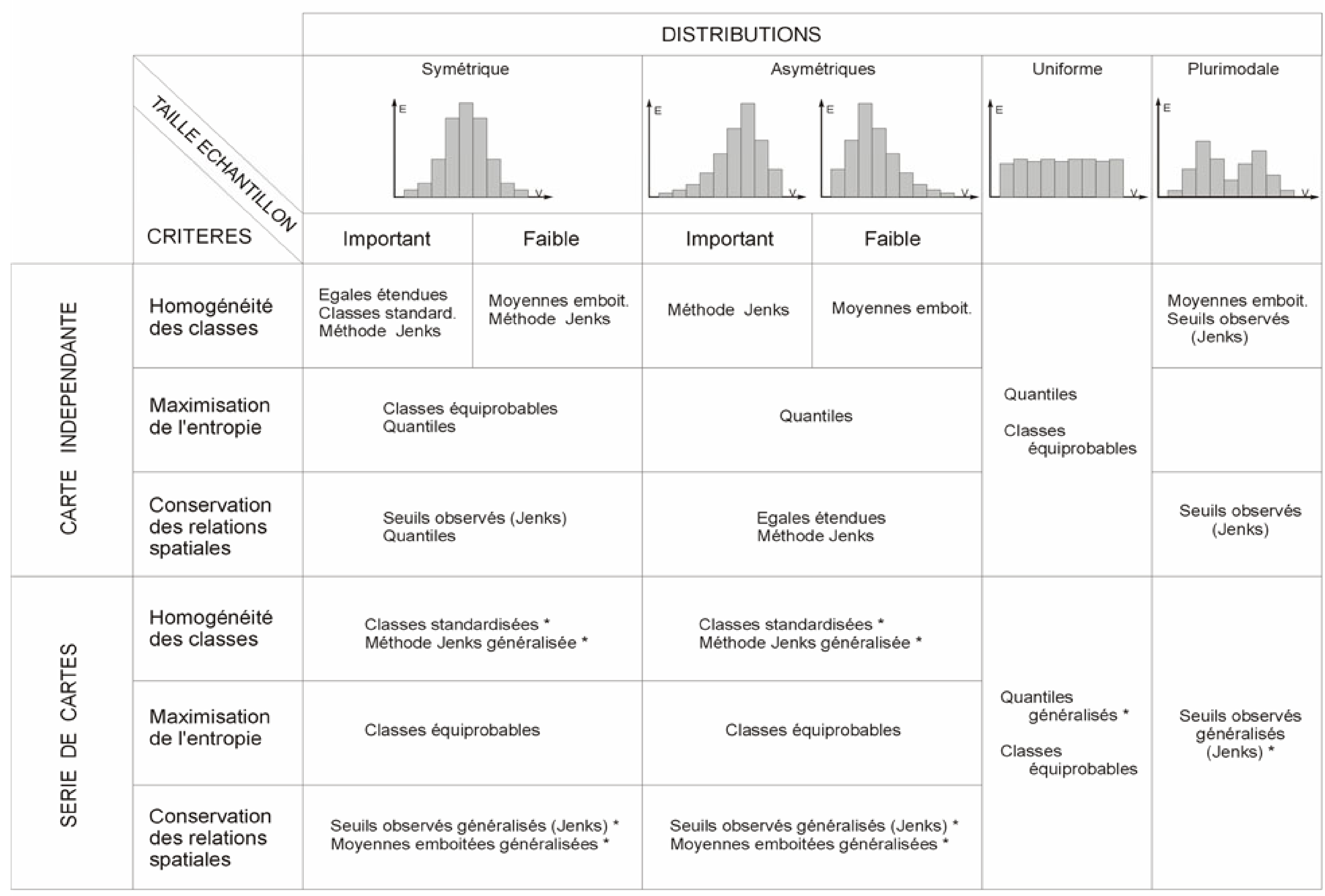
Cf. : Réalisation d'une courbe de fréquences sous Excel.

La répartition des valeurs des individus dans la variable va nous permettre d'avoir une idée synthétique de cette variable. C'est cette "forme" qu'il va falloir essayer de reproduire avec le découpage en classes, si l'on veut préserver au maximum l'information de départ.

Le graphique de fréquences nous donne donc une image de référence de la variable, qui va nous servir de point de comparaison avec les fréquences obtenues après discrétisation de la variable selon les différentes méthodes.

La méthode de discrétisation automatique dont le résultat ressemblera le plus aux fréquences de la variable d'origine sera donc celle qui sera la moins réductrice et la plus conforme.

Guide pour le choix de la méthode automatique (cf. page suivante.)



METHODES ET CHOIX DE DISCRETISATION : tableau indicatif de synthèse

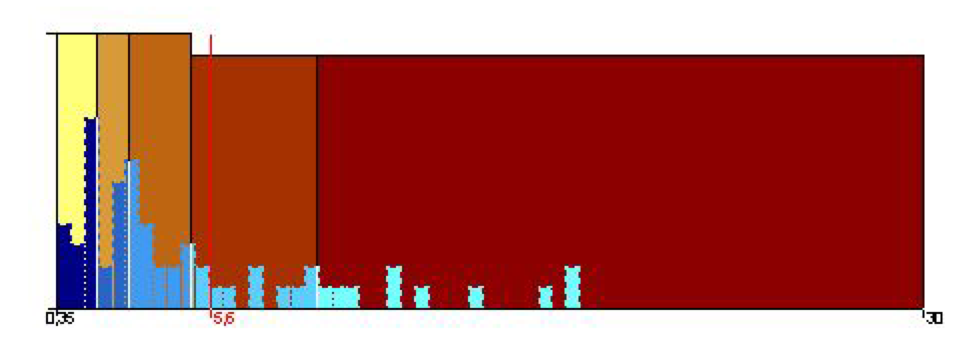
	CRITERES / TAILLE ECHANTILLON	DISTRIBUTIONS : Symétrique – Important	Symétrique – Faible	Asymétriques – Important	Asymétriques – Faible	Uniforme	Plurimodale
CARTE INDEPENDANTE	Homogénéité des classes	Egales étendues Classes standard. Méthode Jenks	Moyennes emboit. Méthode Jenks	Méthode Jenks	Moyennes emboit.		Moyennes emboit. Seuils observés (Jenks)
	Maximisation de l'entropie	Classes équiprobables Quantiles		Quantiles		Quantiles Classes équiprobables	
	Conservation des relations spatiales	Seuils observés (Jenks) Quantiles		Egales étendues Méthode Jenks			Seuils observés (Jenks)
SERIE DE CARTES	Homogénéité des classes	Classes standardisées * Méthode Jenks généralisée *		Classes standardisées * Méthode Jenks généralisée *			
	Maximisation de l'entropie	Classes équiprobables		Classes équiprobables		Quantiles généralisés * Classes équiprobables	Seuils observés généralisés (Jenks) *
	Conservation des relations spatiales	Seuils observés généralisés (Jenks) * Moyennes emboitées généralisées *		Seuils observés généralisés (Jenks) * Moyennes emboitées généralisées *			

* Généralisé : cette expression signifie que l'on prend toutes les valeurs de toutes les variables simultanément

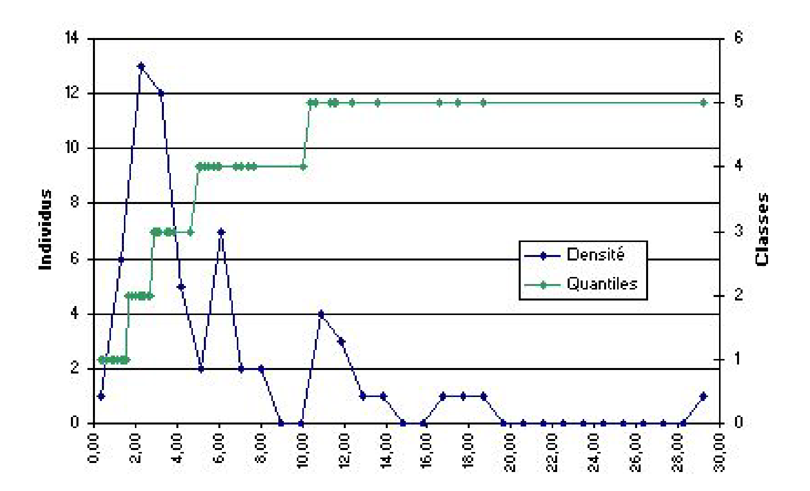
Les différentes méthodes de discrétisation automatique

Là encore, l'utilisation de l'ordinateur va nous permettre de réaliser rapidement des découpages en classes qui seraient autrement long voire impossibles à réaliser. De plus, on va pouvoir comparer très rapidement le résultat de ces discrétisations avec le modèle du graphique des fréquences de la variable.

QUANTILES



(Dans tous les exemples ci-après, le premier graphique présente en dégradé de rouge les classes formées par une discrétisation et en bleu les barres de l'histogramme des fréquences de la variable. Ce graphique est tiré du logiciel SCAP).



(Le second graphique présente en abscisse la variable triée et deux ordonnées : à gauche l'axe du nombre d'individus par barre des fréquences, en bleu, à droite le n° de la classe de chaque individu, en vert).

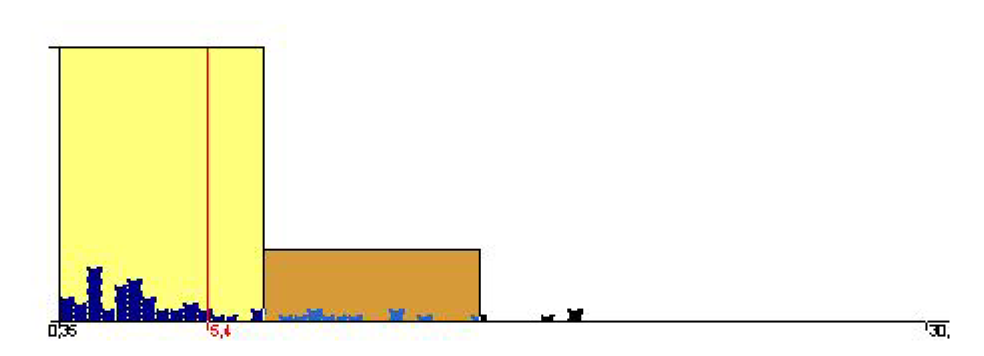
Méthode :
La méthode des quantiles consiste tout simplement à utiliser le même nombre de valeurs par classe, ce qui produit donc des quantiles.

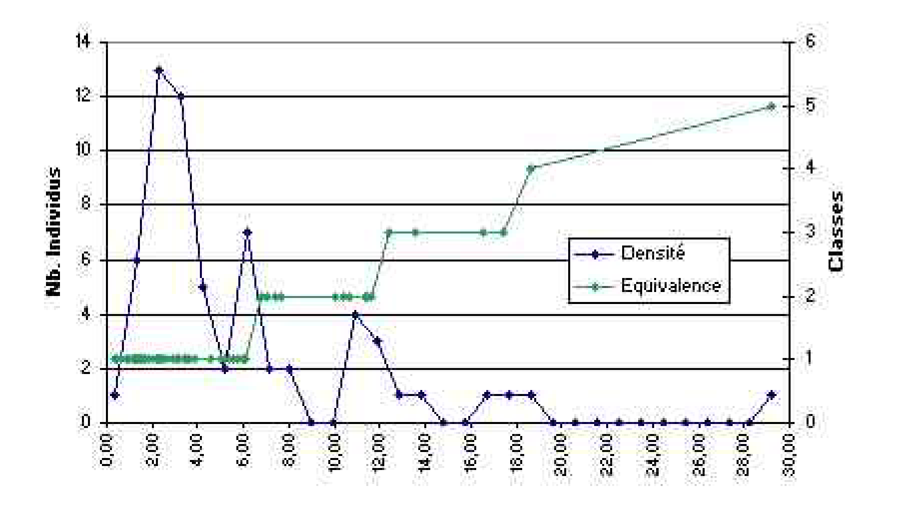
Calcul :
Pour calculer le nombre de valeurs par classe, il suffit de diviser le nombre total de valeurs par le nombre de classes désiré. Ensuite on réalise les classes par comptage.

Objectif :
Cette méthode permet d'obtenir une carte dans laquelle les couleurs de la légende se répartissent à parts égales : donc une carte harmonieuse et facile à lire. Dans le cas de séries de cartes sur le même fond, donc avec le même nombre de valeurs, les classes auront le même poids d'une carte à l'autre, ce qui facilite les comparaisons (chaque classe représente le même pourcentage de la variable totale).

Limites :
Mais cette méthode de discrétisation n'est pas liée à la variable cartographiée elle-même, donc elle ne s'applique correctement qu'aux variables dont le comportement sur le graphique des fréquences n'est pas trop atypique.

ÉQUIVALENCES





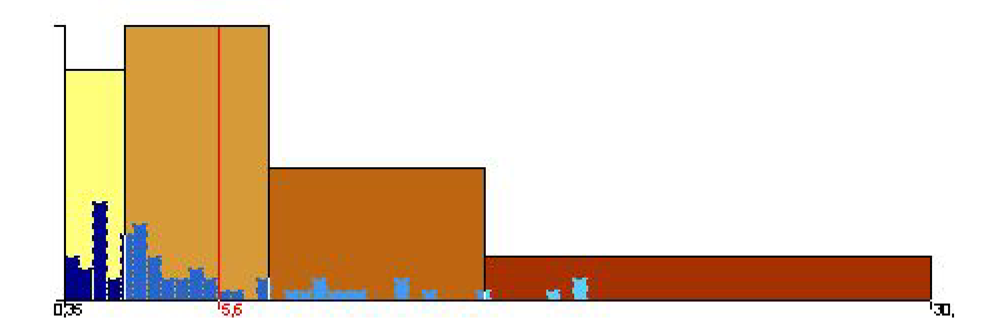
Méthode :
Les équivalences consistent à déterminer des classes de taille égale, du point de vue de l'intervalle de valeurs. Ainsi, chaque classe couvrira une portion identique de l'étendue totale de la variable.

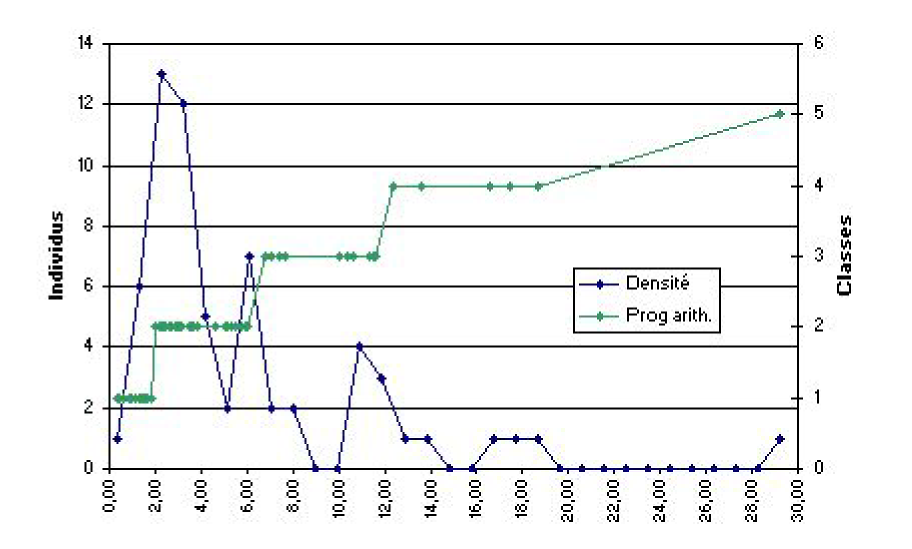
Calcul :
On calcule d'abord l'étendue totale de la variable (maximum-minimum), puis on divise cette étendue par le nombre de classes voulu. La valeur obtenue est l'intervalle de chacune des classes : la première classe ira du minimum à ce minimum plus l'étendue, etc. jusqu'au maximum de la variable.

Objectif :
Cette méthode de discrétisation n'est pas liée à la variable, donc on peut l'utiliser pour comparer des variables distinctes, des phénomènes différents. On peut aussi y voir un moyen de mesure de la distorsion par rapport à une distribution uniforme.

Limites :
Cette méthode est peu utilisée à cause de limites importantes. La plus grave est qu'il arrive souvent qu'une classe ainsi définie ne comprenne aucune valeur, ce qui fausse la discrétisation. Ce problème se pose notamment lorsque l'étendue de la variable est tirée par des valeurs extrêmes, atypiques. Dans des cas moins graves, cette discrétisation à tendance à produire des classes de poids très inégal, et donc des cartes relativement hétérogènes.

PROGRESSION ARITHMÉTIQUE





Méthode :
Cette discrétisation est basée sur une taille croissante de l'intervalle de valeur de chaque classe, en fonction d'une progression d'ordre arithmétique.

Calcul :
La raison de la suite arithmétique des étendues des classes est :

r = (Maximum - Minimum) / (1+2+3+...+nombre de classes)

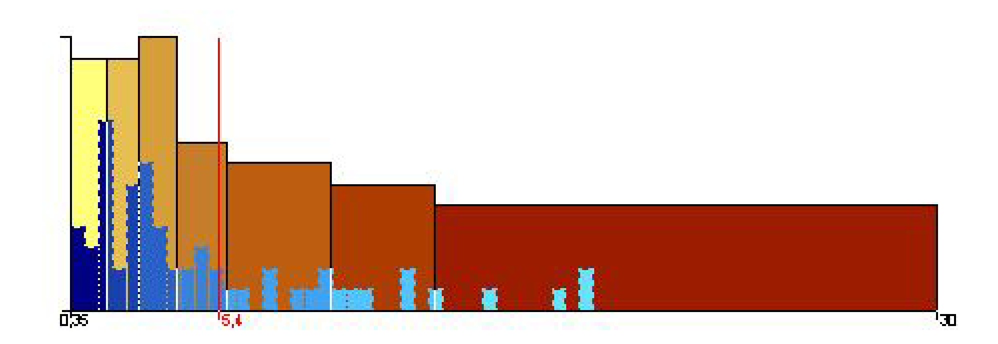
L'étendue de chaque classe est donc calculée comme ceci :

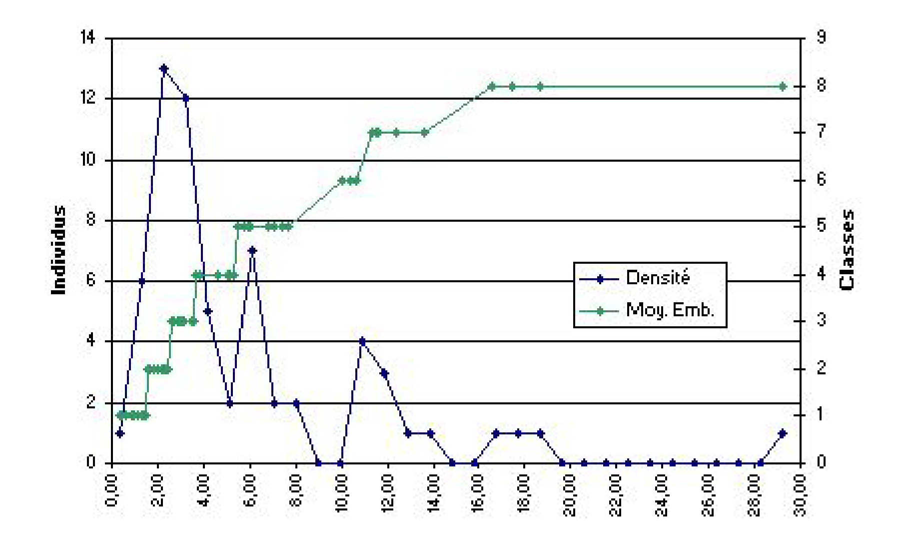
Etendue= r * Numéro de la classe

Objectif :
L'objectif de ces classes de taille croissante est une meilleure prise en compte des distributions fortement dissymétriques à gauche, dans les faibles valeurs de la variable. En effet, au début de l'histogramme des fréquences, où le nombre de valeurs est important, les classes sont de plus petit intervalle, puis leur taille augmente, alors que le nombre de valeurs diminue.

Limites :
Cette méthode ne peut s'appliquer correctement qu'à une forme précise de distribution. Elle est donc très spécifique et ne s'applique avec succès que rarement sur les séries de variables.

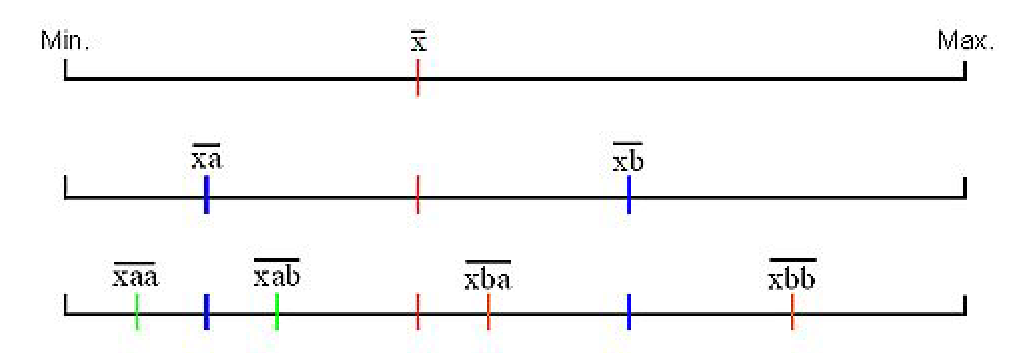
MOYENNES EMBOÎTÉES





Méthode :
La discrétisation par les moyennes emboîtées consiste à découper la variable en utilisant des moyennes successives comme limites de classes.

Calcul :

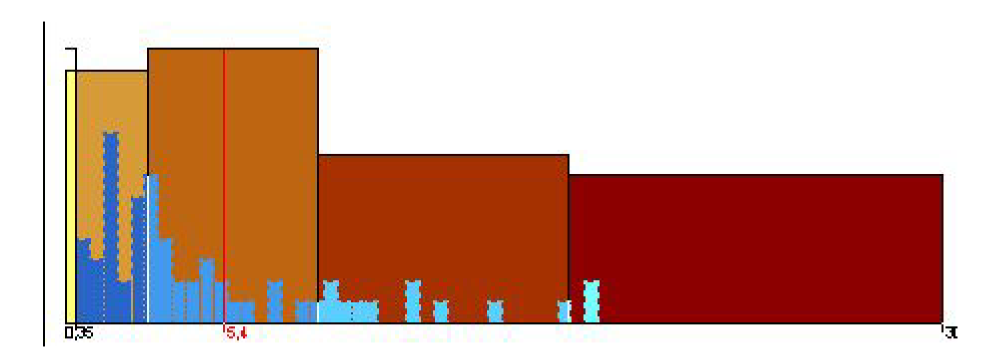


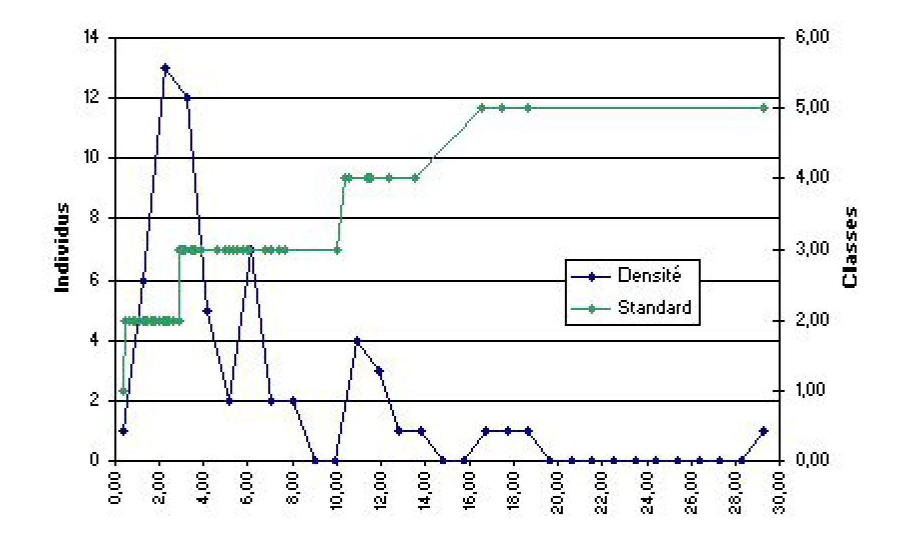
Cette méthode ne peut produire que quatre ou huit classes exploitables. Voilà pourquoi, étape par étape :
1- Calculer la moyenne de la variable, M. Cette moyenne découpe la variable en deux classes, A et B, de part et d'autre.
2- Calculer la moyenne MA de la partie A de la variable. MA découpe la partie A en deux classes : A1 et A2.
3- Calculer la moyenne MB de la partie B de la variable. MB découpe la partie B en deux classes : B1 et B2.
4- On obtient donc quatre classes, A1, A2, B1 et B2, en ayant calculé trois moyennes emboîtées.

Objectif :
Cette méthode a l'avantage d'être fortement liée à la distribution de la variable. Si vous réalisez une carte unique et que vous voulez "coller" aux mouvements de la distribution, cette méthode est généralement assez bien adaptée. Autre avantage : comme les limites des classes sont des moyennes, les cartes produites sont aisément compréhensibles.

Limites :
De par sa méthode de calcul, cette discrétisation peut produire des classes vides ou très hétérogènes, dans le cas de distributions très dissymétriques, voire plurimodales. Enfin, le nombre de classes produit peut ne pas être adapté au nombre d'individus de la variable.

MÉTHODE STANDARD





Méthode :
»Standard deviation » signifie écart-type en anglais. Cette discrétisation est basée sur la moyenne à laquelle on enlève ou on ajoute des parties ou des multiples d'écart-type.

Calcul :
On calcule d'abord la moyenne M de la variable, puis son écart-type E. Ensuite, suivant le nombre de classes voulu, on applique la procédure suivante :

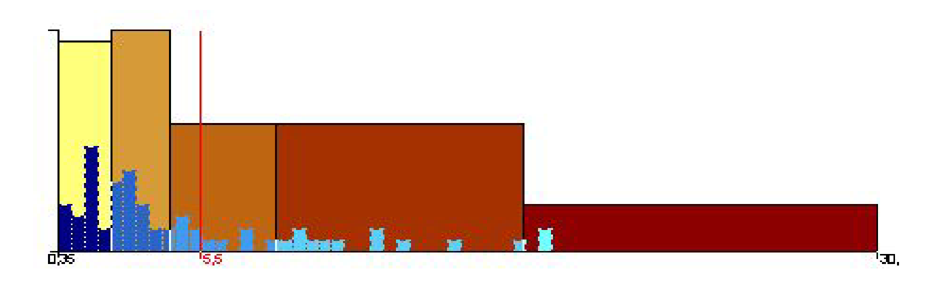
3 classes:
Min, m-0.5E, m+0.5E, Max
5 classes:
Min, m-1.5E, m-0.5E, m+0.5E, m+1.5E, Max

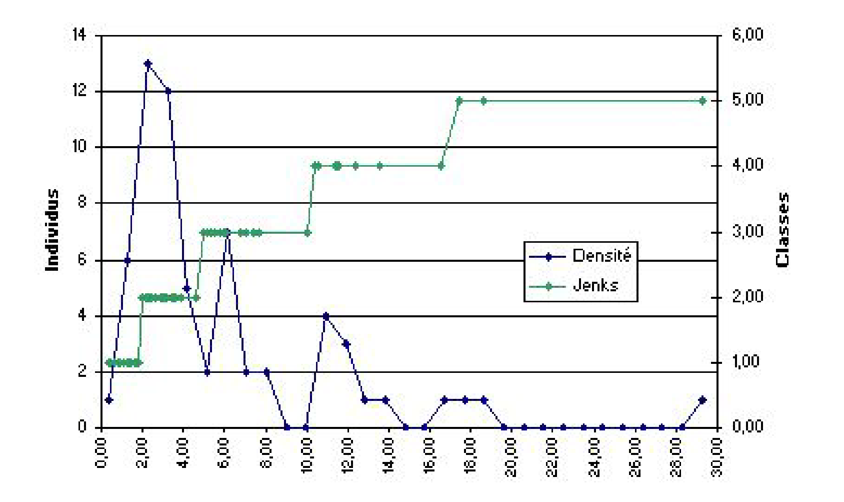
7 classes:
Min, m-2.5E, m-1.5E, m-0.5E, m+0.5E, m+1.5E, m+2.5E, Max

Objectif :
Cette méthode est basée sur des indicateurs statistiques de centralité et de dispersion. Elle permet d'obtenir des séries de cartes comparables entre elles, avec un découpage en classes sensible à la dispersion de la variable. Vous remarquerez que les classes centrales sont plus "courtes" que les classes externes : cette méthode s'applique correctement sur des distributions normales, en courbe de Gauss.

Limites :
C'est encore une méthode très liée à la variable et à sa dispersion. Donc à ne pas utiliser sur des dispersions trop grandes, des dissymétries trop importantes, sous peine de classes vides. De plus, comme pour les moyennes emboîtées, cette méthode s'organise autour d'une moyenne centrale. Donc elle ne fonctionnera pas très bien avec des distributions plurimodales.

MÉTHODE DE JENKS





Méthode :
La méthode de Jenks est l'une des plus performantes, car elle permet d'obtenir des classes très proches des formes de la distribution de la variable, en créant des classes intra homogènes et extra hétérogènes.

Calcul :
Le calcul de cette discrétisation est particulier : en fait SCAP calcule un indice de pertinence pour toutes les combinaisons possibles de classes, suivant le nombre d'individus et le nombre de classes voulu. C'est pourquoi une optimisation mathématique est utilisée : un algorithme convergeant du type nuées dynamiques.

Objectif :
De par sa méthode de calcul, elle est mieux adaptée aux variables multimodales, et possédant un nombre d'individus assez important pour pouvoir faire ressortir des seuils. Cette méthode est un équivalent mathématique assez intéressant de la méthode manuelle des seuils observés.

Limites :
Cette méthode est gourmande en calcul et donc en puissance processeur. Elle fait partie des discrétisations statistiques liées à la variable, donc si elle est très performante sur les cartes uniques, elle n'est pas très utile pour une série de cartes.

RÉALISATION DE CARTES EN PLAGES DE COULEURS

Paramétrage et Cartographie

Conseils généraux de réalisation

Cette série de guides est issue de l'expérience synthétisée dans les excellents ouvrages de Edward R. Tufte (cf. la bibliographie).

- La carte doit montrer la variable, c'est l'information qui doit être perçue en premier lorsqu'on lit la carte.
- La carte doit avoir un objectif clair : description, exploration, mise en forme, illustration.
- La carte doit inciter le lecteur à se poser des questions à propos de la répartition spatiale de la variable, pas à propos de la méthode de représentation utilisée.
- Essayez de réduire au maximum la distorsion de la variable induite par la représentation graphique.
- Essayez d'utiliser le moins de place possible pour présenter le plus d'information.
- La carte doit montrer la cohérence de la variable présentée.
- Essayez de montrer la variable à différents niveaux d'interprétation : globale, locale, structure fine.
- La carte doit être intégrée au document, liée aux textes qui l'entourent.

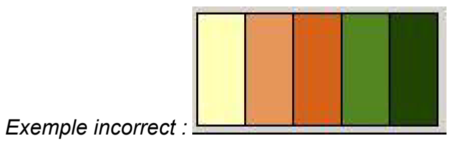
Choix des gammes de couleur

On a vu dans la leçon précédente, sur les discrétisations, comment prendre en compte une bonne partie des contraintes de la variable. On va s'intéresser maintenant aux problèmes posés par l'utilisation de couleurs dans un document graphique.

Le principe d'utilisation de la couleur est de lier une variation d'intensité, de densité d'une teinte à la variation de valeur centrale de la classe.

Ainsi, si on a une discrétisation en équivalence, avec le même écart entre les centres de classes, on peut utiliser le même écart de densité de nuance entre les tons de légende.

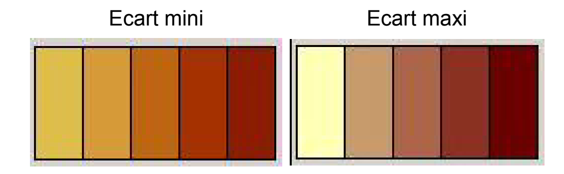
On utilise toujours une teinte par variation détectée dans la variable. Une seule progression de valeur doit être transcrite par une seule progression de densité dans une seule teinte, une seule couleur.

Exemple incorrect : 

En fait, cette liaison, si elle est théoriquement indispensable, est assez difficile à réaliser lorsque l'on a des classes disparates.

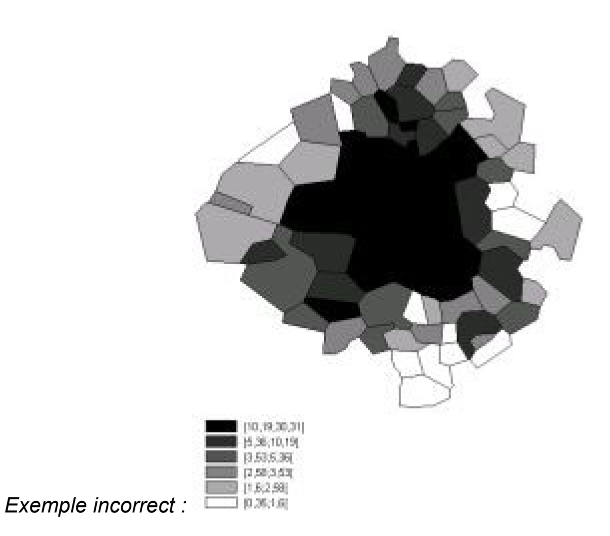
On essaye de maximiser la différenciation de deux tons successifs, pour rendre la lecture de la carte plus simple.

Ecart mini Ecart maxi



En général, on affecte donc la teinte la plus claire possible (mais distinguable sans ambiguïté du fond) à la classe la plus faible, et la teinte la plus foncée, sombre, dense (mais distincte du noir qui est réservé pour le dessin du fond) à la classe la plus forte.

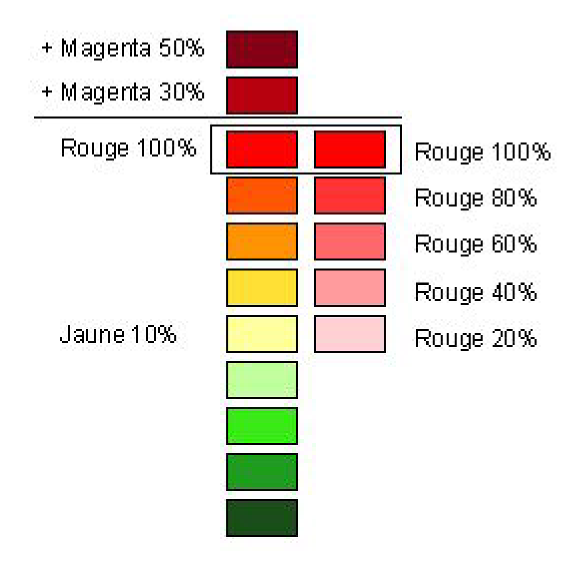
Les couleurs de la variable doivent se distinguer clairement du blanc de la page et du noir du fond de carte et des textes.



Exemple incorrect :

Avec certains logiciels de CCAO il est possible de contourner le problème en changeant la couleur des polygones du fond de carte pour les individus devant recevoir les couleurs les plus sombres de la palette. Ainsi, on peut utiliser des couleurs proches du noir sans que le fond de carte disparaisse.

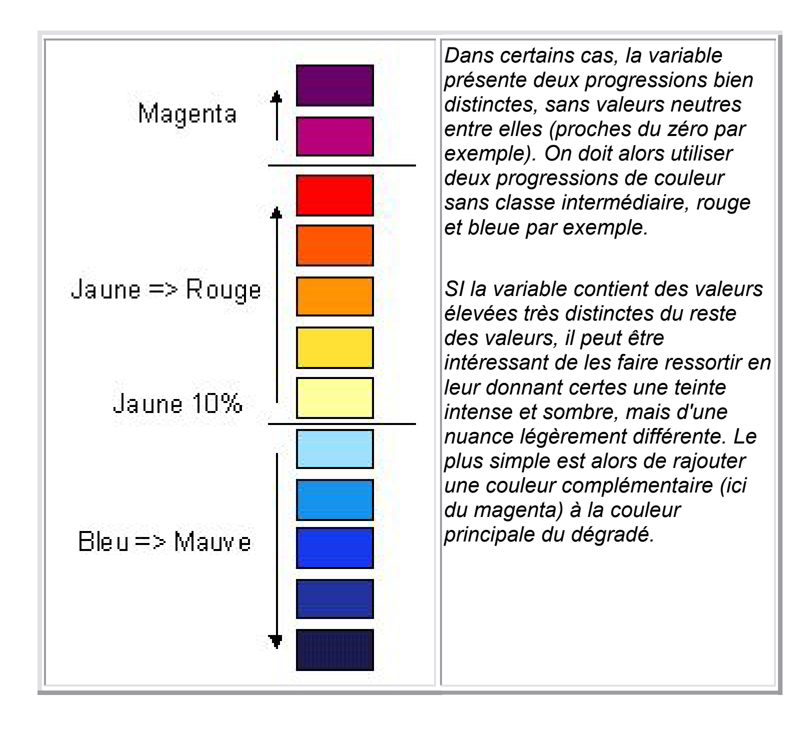
Choix des gammes de couleurs :



Dans la réalisation de palettes en deux progressions, il existe une tradition, une habitude de présentation qui veut que l'on utilise une couleur chaude (rouge, jaune) pour représenter les variations positives et une couleur froide (bleu, vert bleu) pour représenter les négatives.

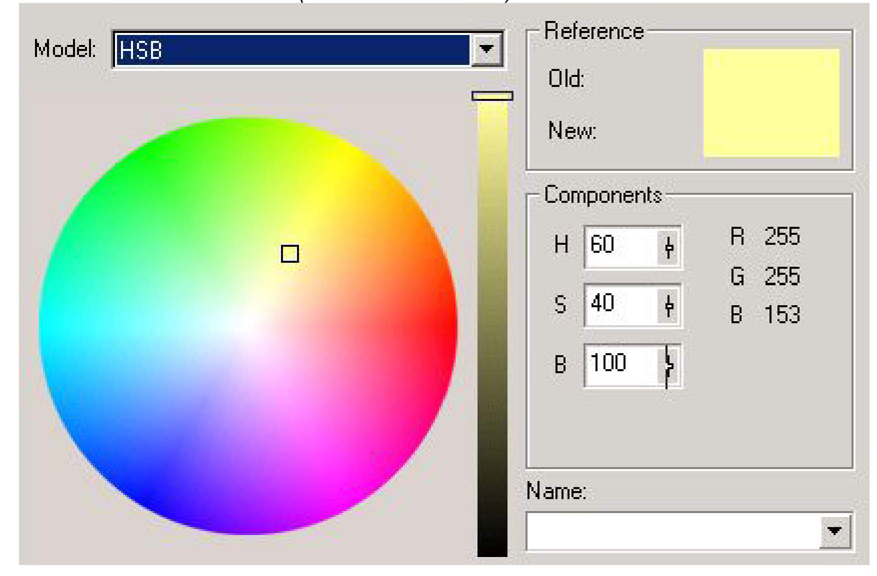
Le jaune est souvent utilisé comme point de départ car il fait bien la liaison entre le rouge et le vert et permet de donner plus de luminosité aux teintes faibles.

Le même principe de complément de la couleur primaire par une couleur secondaire est utilisé lorsque l'on veut renforcer la différenciation des teintes successives lorsque l'on étend la palette de couleurs. Ici on a ajouté deux teintes supplémentaires en ajoutant du magenta au rouge.



Les pourcentages indiqués correspondent à la densité de la teinte principale. Vous pouvez retrouver ces valeurs dans les logiciels de dessin en utilisant une gamme colorimétrique de type TSL : Teinte, Saturation, Luminosité (HSB en anglais).

Mode HSB (CSL basé sur la roue) dans CorelDraw© :



Essayez d'utiliser toujours des couleurs simples et le moins de teintes différentes possibles : cela garantira une meilleure conformité des couleurs à l'impression (ainsi qu'un tarif moins cher chez l'imprimeur).

La contrainte du noir et blanc

Dans de nombreux cas, on ne sait pas exactement quel va être le moyen d'impression utilisé pour la carte, et la méthode de duplication. On doit alors utiliser la gamme la plus passe-partout, celle dont on est quasiment sur qu'elle passera à l'impression et à la photocopie : le noir et blanc tramé.

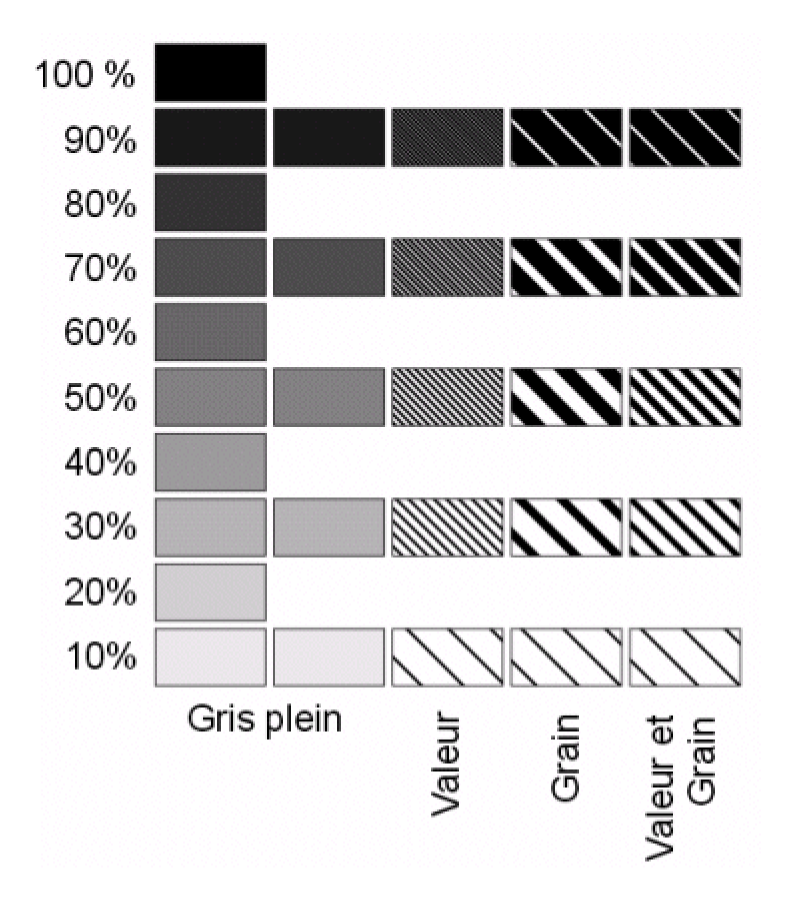
L'utilisation du noir et blanc suppose quelques règles supplémentaires et une contrainte majeure de réalisation, lorsque l'on veut représenter deux progressions.

Un dégradé de gris doit utiliser au maximum l'intervalle entre le gris très clair et le gris très foncé, sans être confondu avec le blanc de la page et le noir du fond. Or les techniques de reproduction les plus répandues ont une fâcheuse tendance à transformer les nuances de gris.

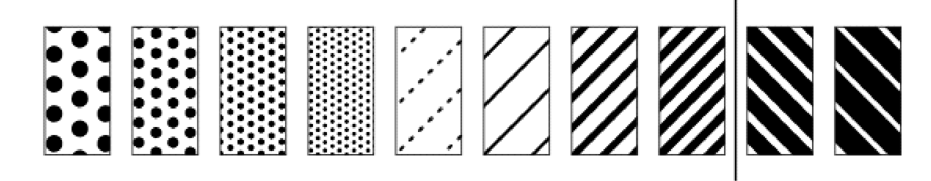
En effet, les techniques d'impression et de reproduction ont des effets différents sur les niveaux de nuance utilisés, et souvent ils ont des limites à ne pas franchir si l'on veut que les dégradés "passent" bien.

N'utilisez pas de gris sous les 10% ni au-dessus des 85%.

Le pire des moyens de reproduction d'une carte, la photocopie, réduit fortement le contraste d'un document, surtout si les nuances de gris ne sont pas tramées, c'est à dire composée de motifs. La sémiologie graphique de Bertin nous apporte des moyens performants de réalisation de trames en noir et blanc, principalement la variation de valeur (% de figurés noirs sur le fond blanc) et de grain (largeur des figurés noirs sur le fond blanc).



Si l'on veut représenter un variable présentant des sous catégories marquées, il faut alors utiliser d'autres méthodes de dessin, voire d'autres figurés ("variables graphiques").



Tracés et aplats

Une autre contrainte de dessin assez méconnue est la largeur minimale du trait. En effet, selon la technique d'impression retenue, la largeur du plus fin trait imprimable varie beaucoup. Le problème est que là encore, les logiciels de CCAO et même les logiciels de dessin nous proposent par défaut des traits souvent trop minces pour l'impression.

Les imprimantes de bureau ont beaucoup évolé et sont capables de performances que peu d'imprimeurs réalisent aux tarifs normaux, sans supplément. Ainsi, un nuancier ou des traits de 0.1mm se verront très bien sur votre imprimante, mais une fois imprimés et/ou photocopiés, c'est la catastrophe.

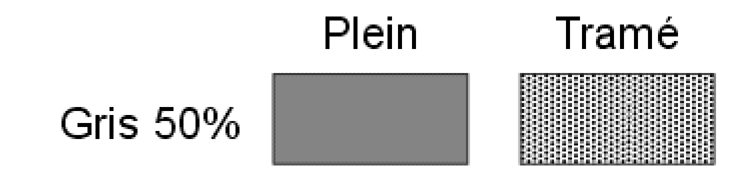

Utilisez des traits d'au minimum 0.2 mm, toujours en noir 100%.

Dans le même domaine, il existe une technique simple pour éviter que les nuances en aplats de gris ne soient transformées par la duplication de votre carte. Les photocopieuses, noir et blanc comme couleur, ont souvent des problèmes pour reproduire les aplats continus sans les maculer de moirés et de taches. Les photocopieurs noir et blanc sont de même optimisés pour le texte (2 couleurs) et pas pour les dégradés de gris.

La solution est simple en théorie, mais parfois difficile à mettre en oeuvre...

Utilisez des aplats tramés et non pleins.

Voici la différence :



A l'œil, à distance normale, la différence est minime car on ne distingue plus les points formant le tramage, mais la photocopieuse, elle, reproduira bien les points, sans faire d'erreurs de contraste puisqu'on reste à deux couleurs : noir et blanc.

RÉALISATION DE CARTES EN PLAGES DE COULEURS

Légende et mise en page

A quoi sert une légende ?

La légende est en fait la clé de lecture d'une carte.

Elle doit donner tous les renseignements nécessaires à la bonne compréhension de l'information de la carte, et ne pas induire en erreur le lecteur par des informations incomplète ou ambiguës.

Une légende doit donc comporter :

- Un titre précisant l'unité dans laquelle est exprimée la variable.
- Une présentation de la liaison classe de la variable / nuance de couleur ou de gris.

Dans tous les cas, les intitulés de la légende doivent reprendre les valeurs numériques composant les bornes des classes de la variable.

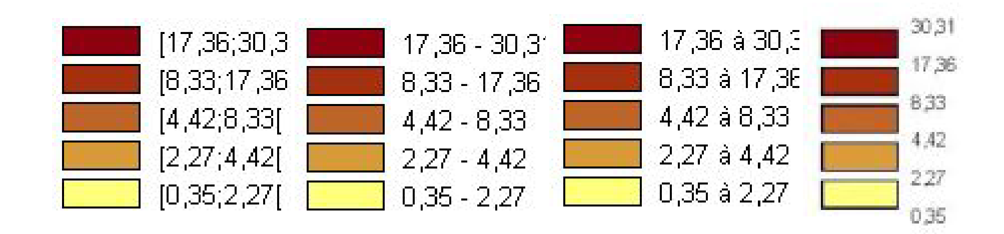
Ce n'est pas la peine de répéter le nom de la variable, le nom de la zone géographique représentée, la date de l'information : c'est le rôle du titre de la carte.

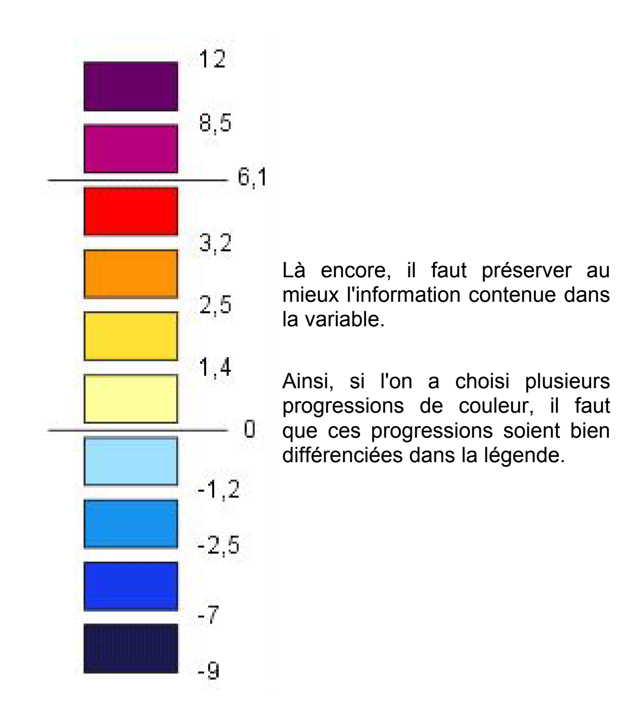
En effet, une carte scientifique se doit d'être simple et claire, donnant toutes les clés de lecture dans le titre et dans la légende, sans commentaires superflus

Présentation de la légende

En ce qui concerne les intitulés des caissons de légende (les rectangles colorés d'une légende de carte choroplèthe s'appellent des caissons), il existes plusieurs solutions de présentation.

On peut présenter ces intitulés de manière mathématiquement juste, mais c'est parfois un peu difficile à lire, alors on peut simplifier :



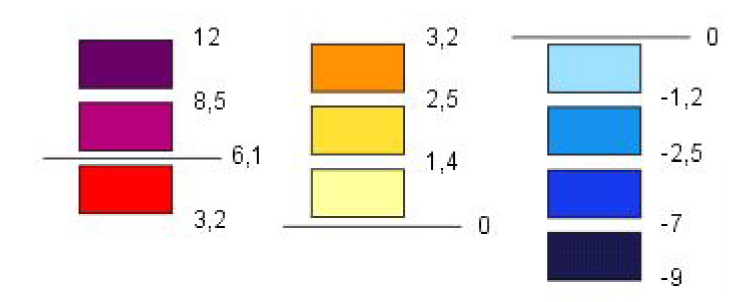


Là encore, il faut préserver au mieux l'information contenue dans la variable.

Ainsi, si l'on a choisi plusieurs progressions de couleur, il faut que ces progressions soient bien différenciées dans la légende.

Une habitude de présentation, devenue tradition, veut que la légende soit toujours présentée avec les valeurs élevées en haut et les valeurs faibles en bas, comme sur un thermomètre.

En cas de manque de place pour la légende, on peut toujours la présenter horizontalement :



Mise en page générale

La mise en page d'une carte comprend certains éléments que l'on placera autour du fond de carte. Mais retenez qu'une carte claire et agréable à lire est une carte simple, sans fioritures.

Une carte doit présenter, autour du fond de carte :

- Un titre qui répond aux questions : Quoi, Où, Quand ?
- Une légende qui renseigne sur l'unité utilisée et le ratio variable/surface du symbole
- Une échelle graphique (cf. exemples ci-dessous)
- Une flèche indiquant l'orientation de la carte si celle-ci n'est pas évidente (nord)
- Une indication de la source de la carte (pour informer sur la validité des données et leur date de mise au point)
- Une indication de l'auteur de la carte si celui-ci n'est pas celui du document global ou si la carte ne fait pas partie d'une série.

AUTRES TYPES DE CARTES
STATISTIQUES AUTOMATIQUES

Avant-propos

Nous allons présenter ici différents types de cartes, qui ont été inventées pour répondre à des situations précises de type de variable, de phénomène, et de type d'implantation géographique.

Ces types de carte sont très rarement offerts dans les logiciels de CCAO, car elles sont beaucoup plus rarement utilisées. Souvenez-vous que l'objectif d'une carte est de présenter une information, pas une technique de représentation, donc si le lecteur est plus intéressé par la manière dont vous avez dessiné l'information que par la visualisation de cette information, votre carte perd de son utilité.

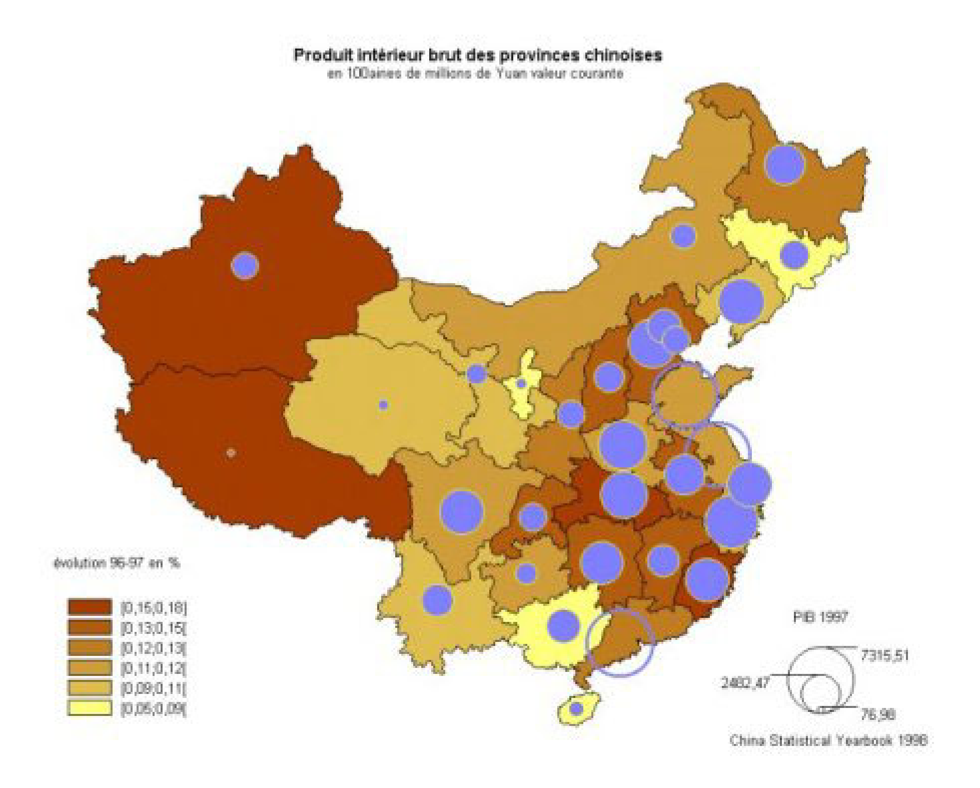
Les cartes doubles

Les cartes dites doubles sont en fait des cartes présentant deux variables sur le même fond de carte.

Elles se présentent sous deux formes :

- Une symbolique par-dessus une carte en plages de couleurs
- Une carte en couleurs à l'intérieur des symboles d'une symbolique

Elle présuppose naturellement que les deux phénomènes représentés ensemble sont complémentaires, voire liés. C'est un type de carte particulièrement adapté lorsque l'on veut représenter l'aspect absolu ET l'aspect relatif d'une variable, par exemple la variation de population absolue entre deux dates et la part que cette variation représente par rapport à la période (taux d'évolution).



Mais attention, si ces cartes sont intéressantes pour gagner de la place et permettre des comparaisons rapides, elles cumulent les contraintes des deux types de cartes. Le problème principal est de préserver la visibilité de chaque représentation partout, donc de ne pas cacher la choroplèthe avec les symboles ou de ne pas utiliser de symboles trop petits pour qu'on puisse en voir la couleur.

Les cartes continues ou lissées

Les cartes lissées ont pour objectif principal de proposer une visualisation continue dans l'espace d'un phénomène alors que l'on n'en possède que des éléments ponctuels.

Cela sous-entend naturellement que l'on sache par avance que le phénomène à représenter est continu dans l'espace (qu'il possède une valeur distincte en tout point de l'espace), mais que l'on a pu faire qu'une quantité limitée de relevés sur le terrain.

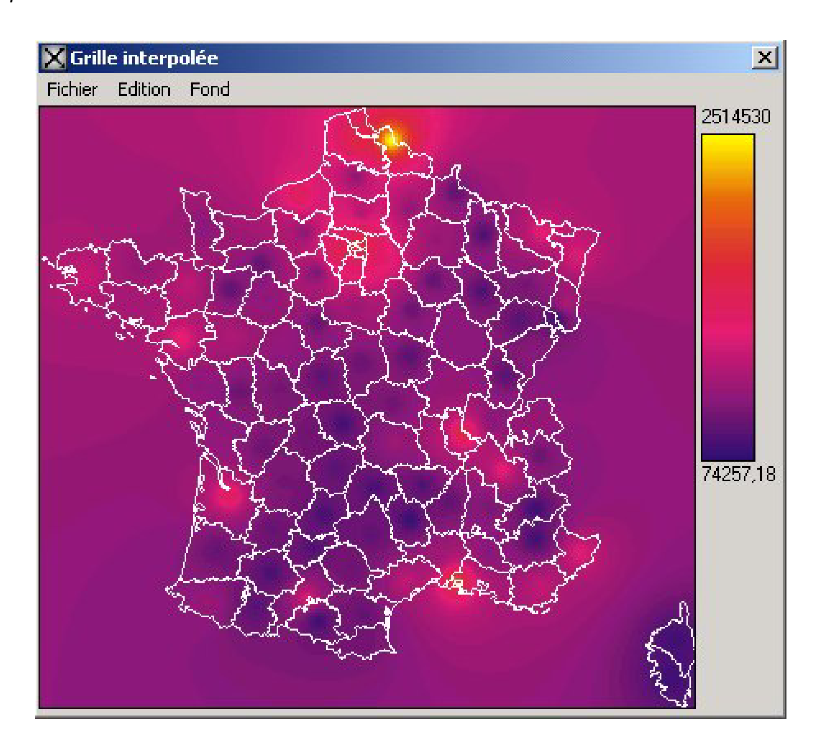
C'est le type de carte intéressant pour représenter les phénomènes naturels généraux, comme en météorologie, en géologie, en biogéographie...

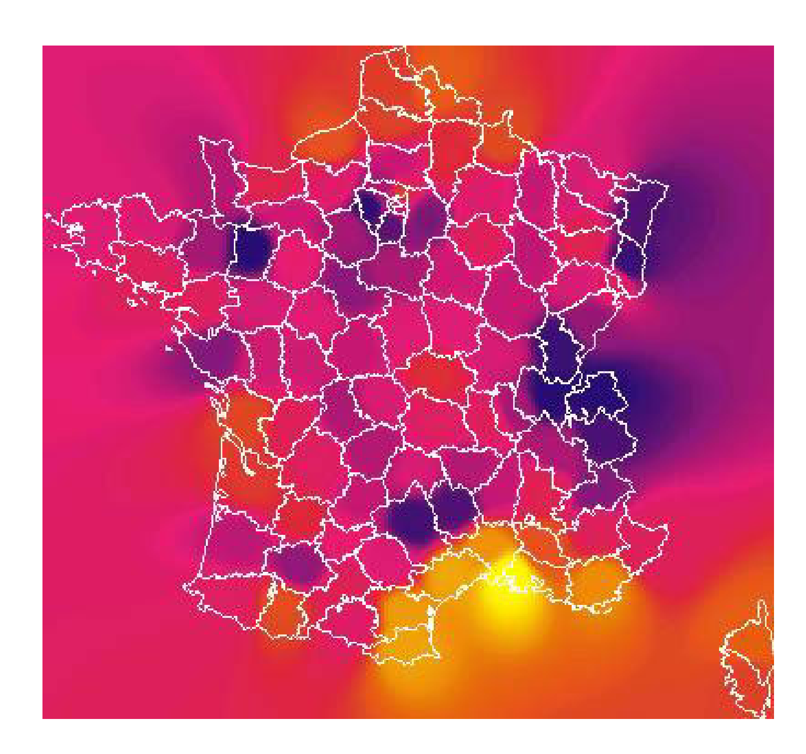
Par exemple, lorsque l'on veut cartographier les variations de température sur une zone, alors que l'on ne possède que les valeurs des stations météorologiques implantées dans cette zone, on est obligés de recourir à ce type de carte.

Pour pouvoir dessiner des plages de couleurs continues, ou des courbes concentriques du type courbes de niveau, il faut que le logiciel utilisé interpole les valeurs de la variable aux endroits où il n'y a pas de valeur connue. Cette création de valeurs est faite le plus souvent par interpolation, en utilisant des méthodes mathématiques plus ou moins complexes, sur une grille plus ou moins fine de points encadrant les localisations des valeurs connues.

La méthode la plus simple consiste, pour un point donné de la grille d'interpolation, à faire une moyenne des valeurs des points connus, en pondérant cette moyenne par le carré de 'inverse de la distance entre chaque point connu et le point interpolé. Cette méthode, à cause de la formule employée, s'appelle la méthode gravitaire.

Exemples :



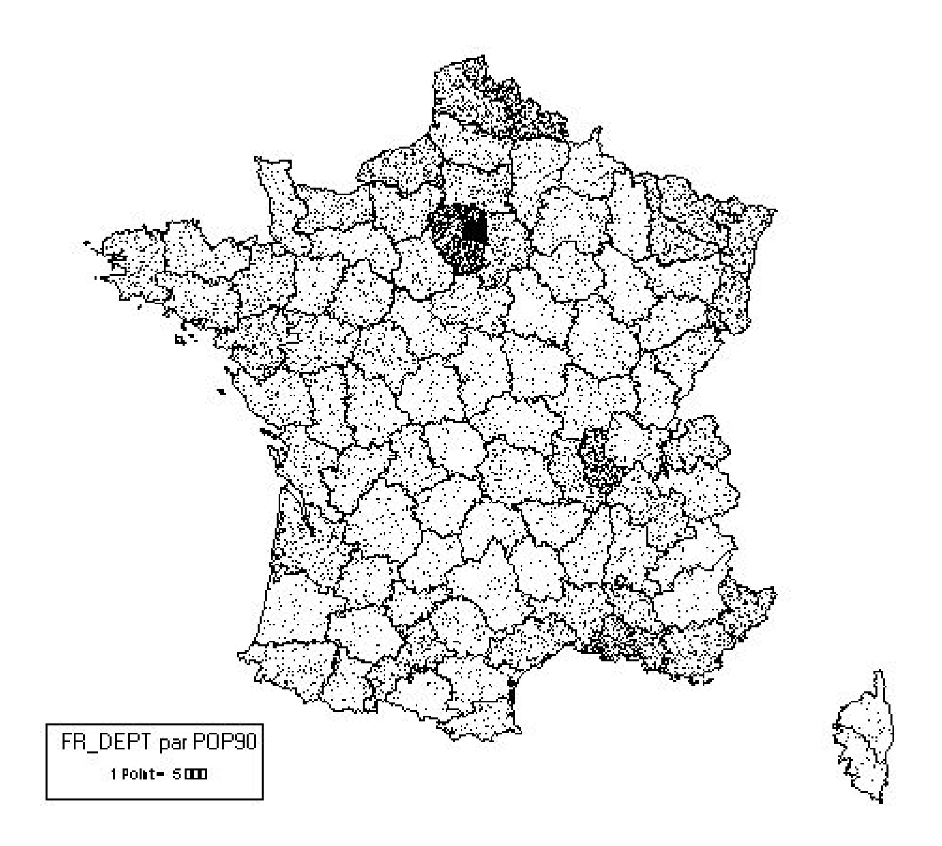


Il existe d'autres méthodes, faisant intervenir les notions de probabilités et de statistiques, qui ont des résultats plus intéressants.

Mais il faut garder à l'esprit que ces cartes présentent des données interpolées, donc fausses (non réelles) par définition, et que la méthode choisie influence fortement sur la valeur prise par un point précis. Donc c'est à la fois créer de l'information et la simplifier pour la représenter, ce qui nous éloigne un peu des valeurs réelles...

Les cartes en densités de points

Ce type de carte présente l'information sous la forme de semis de points réguliers ou aléatoires, dont la densité (le nombre par unité de surface) traduit la valeur de la variable dans le polygone observé.



On utilise cette méthode pour des variables absolues, car on définit une valeur pour chaque point. C'est assez paradoxal de représenter une variable absolue sur une surface, mais en fait ce sont des points qui sont répartis sur cette surface.

Le problème majeur de ce type de carte est que l'on a tendance à inconsciemment associer la localisation d'un point avec celle de la variable, puisque c'est une carte. Or, dans ce cas précis, la localisation d'un point n'a pas de sens, elle est choisie au hasard sur une surface donnée, pour obtenir une dispersion maximale.

Dans les zones où la densité de points est faible, on se retrouve donc avec peu de points, et là la tentation est grande de donner un sens à la localisation de chaque point, comme on a l'habitude de voir les cartes en symboles proportionnels.

Exemple de carte en densités de points peu interprétable :

